



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Wireless multi-carrier systems

Resource allocation, scheduling and relaying

Kaneko, Megumi

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Kaneko, M. (2008). *Wireless multi-carrier systems: Resource allocation, scheduling and relaying*. Center for TeleInfrastructure, Aalborg University.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Wireless Multi–Carrier Systems: Resource Allocation, Scheduling and Relaying

by

Megumi Kaneko

A Dissertation for the Degree of Doctor of Philosophy

defended on Wednesday November 21st, 2007

at Aalborg University

Aalborg, Denmark

Supervisors:

Prof. Ramjee Prasad, Aalborg University, Denmark

Assoc. Prof. Elisabeth de Carvalho, Aalborg University, Denmark

Assist. Prof. Petar Popovski, Aalborg University, Denmark

The assessment committee:

Prof. Søren Holdt Jensen, Aalborg University, Denmark (Chairman)

Prof. Fumiyuki Adachi, Tohoku University, Japan

Prof. Hermann Rohling, Technische Universität Hamburg-Harburg, Germany

Moderator:

Prof. Gert Frølund Pedersen, Aalborg University, Denmark

ISSN 0908-1224

ISBN 87-92078-33-8

Copyright © 2007 by Megumi Kaneko

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Abstract

One of the biggest challenges for the next generation wireless system is the ubiquitous access to high data rate services. At the physical layer, Multi-Carrier (MC) transmission technology is envisaged, due to its high spectral efficiency and robustness to frequency selective fading. While MC transmission opens the field for enabling such ambitious goals, the available radio resource needs to be efficiently distributed in order to perceive the gains at the system level. There are many design issues to be addressed. Two types of system architecture, the cellular and the relayed system, envisioned for the next generation wireless system, are considered in this thesis. For each system, the main target is to produce radio resource allocation and scheduling algorithms that provide good performance with low complexity, making them suitable for practical implementation. While the common scheduling algorithms optimize one given metric, in reality, optimization of different types of metrics is desirable. The objective of the proposed algorithms is to enhance the fairness among users and reduce service delays, without sacrificing system throughput. Another issue in MC systems is the explosion of the amount of Channel State Information (CSI) to report to the scheduling node. The higher the amount of CSI, the better the scheduling performance but the larger the amount of signalling. Breaking with the usual approach where CSI minimization is viewed as a stand-alone problem, independent of the type of algorithm, in some of the proposed algorithms CSI reduction is included as one of the design criteria. Adaptive CSI reduction schemes, which can be used for different types of algorithms, are also developed. The thesis is divided into two parts, each dealing with the considered system architectures.

Allocation algorithms are developed for cellular MC system, with a particular attention towards Proportional Fair Scheduling (PFS). While optimal PFS in the MC case is prohibitively complex, the proposed method provides extremely tight bounds with reduced complexity. Several suboptimal algorithms are developed. As the common algorithms do, the first set of algorithms operate on a slot-by-slot basis, where a slot is equal to a symbol duration. However, the practical radio frames such as in the standards contain a multitude of time slots for which the scheduling decision is made. The second set of algorithms enable multi-slot

scheduling, and take advantage of the additional time dimension by allocating different users in one subchannel—defined in this thesis as a group of adjacent subcarriers—, over different time slots, which is referred to as *User Multiplexing*. The advantage of these algorithms is that they are designed in such a way that reduces the amount of signalling required for describing the user mapping within the frame, whereas the sequential application of slot-by-slot scheduling to all the slots of the frame results in an unacceptably high amount of signalling, dropping the overall throughput. Results show that the proposed algorithms achieve excellent throughput/fairness trade-off and reduce service delays. Moreover, CSI feedback schemes are proposed, characterized by their flexibility to adapt to the required CSI which varies depending on the scheduler. The developed *Differential encoding* mechanism achieves large feedback reductions, even for PFS for which higher amounts of CSI are required to ensure an acceptable level of fairness.

The second type of system is the relay-based MC system. The multitude of parameters in a multi-carrier system with multiple users and relays make the allocation problem very complex. The proposed algorithms require low complexity and minimal CSI feedback as the allocation of relayed users is decoupled from the allocation of BS-originated transmissions. Higher throughput is achieved by the proposed request mechanism, where the relay requests the BS to forward new packets for the users likely to be scheduled over the following frames, thereby optimizing the resource allocation between relayed and non-relayed users. Results show that the algorithms achieve good throughput/fairness trade-off. In the case of multiple relays within a cell, the proposed adaptive schemes provide significant performance enhancement over static schemes. Finally, the issue of cooperative diversity in MC systems is addressed. Approximations of the outage probabilities of designed schemes are theoretically derived. The results show the very high accuracy of the bounds. Simulations with practical OFDM channels confirmed the analysis, and one of the schemes, the *Average Best Relay Selection Scheme* (Avg-BRS), is the most promising due to its near-optimal performance while reducing the required CSI.

Dansk Resumé

En af de største udfordringer for den næste generations trådløse kommunikationssystemer er kravet om adgang til højhastigheds dataservices på ethvert tidspunkt, ethvert sted. Det er tænkt således at på det fysiske lag at multi bærebølge transmissions teknologi er anvendt på grund af dens høje spektrum effektivitet og robusthed mod frekvensafhængig dæmpning. Mens multi bærebølge transmission åbner for mulighederne for at opnå sådanne ambitiøse mål, er det nødvendigt at distribuere de tilgængelige radio ressourcer, for at kunne forstå virkningen på system niveau. Der er i det sammenhæng mange design problemer der skal løses. To typer af system arkitekturer; de cellulære og viderebringende systemer, som forestilles anvendt i næste generations trådløse netværk, bliver undersøgt i denne afhandling. For hvert system er hovedformålet at nå frem til radio ressource allokerings- og planlægningsalgoritmer som giver en god ydeevne med en lav kompleksitet, som gør dem egnede til praktisk implementering. Mens den fælles planlægningsalgoritme optimerer en parameter, er det i virkeligheden ønskværdigt at kunne optimere flere forskellige typer af parameter. Formålet med de foreslåede algoritmer er at forbedre ligestillingen mellem brugere og reducere service forsinkelser, uden at det går ud over system gennemgangen. Et andet problem med multi bærebølge systemer er den nærmest eksploderende mængde af Kanal TilstandsInformation (KTI, eller CSI på engelsk) der skal reporteres til den planlæggende node. Jo flere CSI informationer der er, jo bedre bliver planlægningsydeevnen men jo flere signaler skal der også til. Ved at bryde med den gængse tilgang hvor CSI minimering er set som et alenestående optimeringsproblem, er det i nogle af de foreslåede algoritmer, foreslået CSI reduktion som et design kriterium. Afhandlingen er delt ind i to dele, hver omhandlende de nævnte system arkitekturer.

Allokeringsalgoritmer er udviklet til cellulære multi bærebølge systemer, med især hensyn til proportionel fair planlægning (PFS). Mens optimal PFS i multi bærebølge scenariet er utrolig komplekst, udbyder den foreslåede metode ekstremt tætte grænser med reduceret kompleksitet. Flere, ikke optimale, algoritmer er udviklet herunder. Ligesom den almindelige algoritme gør, opererer de første sæt algoritmer på en per slot basis, hvor et slot er lig med et symbols tidsvarighed. Imidlertid indeholder de reelle radio rammer såsom dem i standarderne, en lang

række tidsslots hvori planlægningsbeslutninger er taget. Det andet sæt af algoritmer sikrer en multislots planlægning og drager fordel af den ekstra tidsdimension ved at allokere forskellige brugere på en delkanal - defineret i denne afhandling som en gruppe af tætliggende delbærerbølger over forskellige tidsslots, som refereres til som bruger multiplexing. Fordelen ved disse algoritmer er at de er designede på sådan en måde at de reducerer mængden af signallering der ellers er påkrævet for at beskrive bruger afbildningen indenfor en ramme, til forskel fra den sekventielle applikation af et per slot planlægning af alle slots i rammerne, som resulterer i en uacceptabelt høj mængde af signallering, hvormed den gennemsnitlige data gennemgang mindskes. Resultaterne viser at den foreslåede algoritme opnår et glimrende forhold mellem data gennemgang og retfærdighed, og reducerer service tidsforsinkelser. Derudover, er CSI feedback metoder foreslået, som er karakteriserede ved deres fleksibilitet ved tilpasning af den påkrævede CSI som varierer afhængigt af planlæggeren. Den udviklede Differentiel Indkoder mekanisme opnår en stor feedback reduktion, selv for PFS hvori en høj mængde af CSI ellers er påkrævet til at sikre et acceptable niveau af retfærdighed.

Den anden type af systemer er de viderebringende baserede multi bærebølge systemer. Mængden af parametre i et multi bærebølge system med flere brugere og viderebringere gør reserverings problemet meget komplekst. Den foreslåede algoritme kræver lav kompleksitet og et minimalt CSI feedback så reserveringen af viderebragte brugere er afkoblet fra reserveringen af transmission kommende fra base stationer. En højere data gennemgang er opnået ved den foreslåede forespørgselsmekanisme, hvor viderebringeren forespørger base stationen om at sende den nye data pakke videre for brugere der højst sandsynligvis bliver planlagt indenfor de efterfølgende data rammer, hvilket optimerer ressource reserveringen mellem den viderebringende og de reelle brugere. Resultaterne af dette viser at den foreslåede algoritme opnår et godt forhold mellem god data gennemgang og retfærdighed mellem brugere. I scenariet med flere viderebringere indenfor en celle, opnår algoritmen en signifikant ydelses forøgelse i forhold til en statisk tilgang. Endelig er problemet vedrørende samarbejdende forskellighed i multi bærebølge systemer også behandlet. Tilnærmelse af udfaldssandsynligheder af den designede metode er teoretisk udledt. Resultaterne viser den meget høje nøjagtighed af grænserne. Simulationer med praktiske antagelser bekræfter analysen, og metoden ved navn Average Best Relay Selection Scheme (AvgBRS) er den metode der ser mest lovende ud på grund af dens nær optimale ydeevne samtidig med at den reducerer den krævede Kanal TilstandsInformation (CSI).

Acknowledgements

This thesis is the achievement of a long and sinuous journey, which begun way before starting my Ph. D. study. The journey itself would not have been attempted without the help and encouragements of all the people who surrounded me, in each stage of my life.

First of all, I would like to thank my principal supervisor, Prof. Ramjee Prasad, for giving me the opportunity to work for this Ph. D., and for always providing me with his support during these three years. I would like to thank my co-supervisors, Assoc. Prof. Elisabeth de Carvalho and Assist. Prof. Petar Popovski, for their continuous support and interest to my Ph. D. study. I address my profound gratitude to Assist. Prof. Petar Popovski, with whom I have worked most closely during these three years. There is no doubt that his great scientific and academic sense combined with his exceptional guidance and availability, even during the toughest times, helped me to identify and solve the key problems which build this thesis.

The first two years of this work were devoted to the JADE project, run between Aalborg University and Samsung Electronics, Korea. It was an immensely valuable experience to work with the industry, so I would like to thank Samsung Electronics for supporting this work. It was a pleasure to be part of the JADE team, and I thank all my colleagues from Samsung and from Aalborg University.

I express my deep gratitude to Prof. Hideaki Sakai, Assoc. Prof. Kazushi Ikeda and Assist. Prof. Kazunori Hayashi for accepting me during six months within their research team in Kyoto University, Japan. I also would like to thank all my supervisors for their approval of this period abroad. I am truly obliged to the International Doctoral School of Aalborg University for pushing me to plan this study abroad and for their grant which concretized it. I would like to address my special thanks to Assist. Prof. Kazunori Hayashi for providing me with his unfailing guidance and feedback during this period. This collaboration with Kyoto University was very positive and fruitful, as the results obtained constitute the last part of the thesis.

In addition to my supervisors, I address my extreme gratitude to Assoc. Prof. Hiroyuki Yomo and Assist. Prof. Kazunori Hayashi for reading my thesis. Their

ACKNOWLEDGEMENTS

invaluable comments and advices improved tremendously the quality of this thesis.

I am deeply indebted to the Georges Besse Foundation, France, which awarded me with their grant for pursuing my engineering studies in France. Without their support, I would not have arrived here. I also would like to thank my French engineering school, Institut National des Télécommunications, France, for giving me the opportunity to pursue my MSc. studies in Aalborg University.

The whole Ph. D. process would not have been possible, or at least enjoyable, without the everyday support from all my colleagues. Many thanks to all my colleagues from the former WING group, and to all my present colleagues in the APNet section. Special thanks to the colleagues and students whom I met in Kyoto University, for their kindness and for offering me their help whenever I needed.

My deepest thanks to all my friends in Denmark, Japan and France for their care and support. One of the greatest gifts in coming to Aalborg was to be linked to so many people, who come from so different places around the world.

I address my profound gratitude to my parents and all my family in Japan. I am truly grateful for their trust and love, which have carried me all along. Finally, I am immensely grateful to Hiro, who, whether in the good or bad times, was always by my side. Every day, he has given me motivation and strength. I dedicate this thesis to my parents and to Hiro, who have unconditionally supported and encouraged me all these years.

Contents

Abstract	iii
Dansk Resumé	v
Acknowledgements	vii
Acronyms	xv
1 Introduction	1
1.1 Motivation	1
1.2 The Problems Defined	2
1.2.1 Radio Resource Allocation and Scheduling in a Multi-Carrier (MC) Cellular System	4
1.2.2 Radio Resource Allocation and Scheduling in a MC Relay System	5
1.3 Thesis Contributions	6
1.4 Thesis Outline	7
2 Background	9
2.1 Introduction	9
2.2 System Architecture	9
2.3 Multi-User Diversity	11
2.4 Multi-Carrier System	11
2.5 Proportional Fair Scheduling	12
2.6 Channel State Information	14
2.7 Relay System: Conventional Relaying and Cooperative Diversity .	15
2.8 System Assumptions	17
2.8.1 Assumption on Adaptive Modulation and Coding (AMC) .	17
2.8.2 Assumptions on Interference	18
2.8.3 Assumptions on Power Allocation	18
2.9 Related Work	18

CONTENTS

2.10	Summary	20
I	Radio Resource Allocation and Scheduling in Cellular Multi-Carrier System	21
3	Proportional Fairness in a Multi-Carrier System	23
3.1	Introduction and motivation	23
3.2	Bounds of the Optimal MC-PF Metric in the Single-Slot case . .	24
3.3	Upper Bound of the Optimal MC-PF Metric	24
3.3.1	Lower Bound of the Optimal PF Metric by Rounding the Solution to a Discrete Feasible Point	27
3.4	Proposed Approximation Algorithms	27
3.4.1	Trade-off between Rate and Fairness: Best-Rate and Best- PFS algorithms	27
3.5	Numerical Results	30
3.6	Summary	32
4	Proportional Fairness in a Multi-Carrier System with Multi-Slot Frames	33
4.1	Introduction and Motivation	33
4.2	Proportional Fair Scheduling for Multi-Carrier System, in the Multi- Slot case	35
4.3	Bounds of the Optimal MC-PF Metric in the Multi-Slot case . . .	36
4.3.1	Upper Bound of the Optimal-PF Metric in the Multi-Slot case	36
4.3.2	Lower Bound of the Optimal PF Metric by Rounding the Solution to a Discrete Feasible Point in the Multi-Slot case	37
4.4	Proposed Algorithms with User Multiplexing	38
4.4.1	PF-User Multiplexing Algorithm (PF-Mux)	39
4.4.2	Rate-User Multiplexing Algorithm (Rate-Mux)	39
4.4.3	Rule for Definite Subchannel Ordering	40
4.4.4	Specification of a Reference Algorithm	40
4.5	Allocation overhead	40
4.6	Numerical Results	43
4.7	Summary	49
5	Adaptive Provision of CSI Feedback in OFDMA Systems	51
5.1	Introduction	51
5.2	System Model	52
5.3	Amount of Useful Feedback	53
5.3.1	Dependency on the Number of Users	53

CONTENTS

5.3.2	Dependency on the Scheduler	53
5.4	CSI Feedback Encoding Schemes	55
5.4.1	Reference Scheme	55
5.4.2	Basic Schemes for Encoding one AMC Level	55
5.4.3	Proposed Grouped Encoding Scheme	56
5.4.4	Proposed Differential Encoding Scheme	56
5.4.5	Complete Process for Feedback Encoding	58
5.5	Numerical results	58
5.6	Summary	62

II Radio Resource Allocation and Scheduling in Relay System 63

6	Throughput–Guaranteed Resource Allocation Algorithms for Relay–aided Cellular OFDMA System	65
6.1	Introduction and Motivation	65
6.2	System Model	66
6.2.1	Frame structure	66
6.2.2	Path Selection	68
6.3	Proposed Resource Allocation Algorithms for the Case of Single Relay	71
6.3.1	Fixed Time Division (FTD) Algorithm	72
6.3.2	Adaptive Time Division (ATD) Algorithm	74
6.3.3	Discussion on the packet requests to the BS by the relay	76
6.4	Proposed Resource Allocation Algorithms for the Case of Multiple Relays	77
6.4.1	Multiple–RS Parallel (MRP) Algorithm	77
6.4.2	Multiple–RS Parallel with Activation (MRPA) Algorithm	77
6.4.3	Multiple–RS Adaptive Activation (MRAA) Algorithm	78
6.5	Design of Optimized Algorithms for Performance Assessment	79
6.5.1	Upper Bound for Throughput in the case of Single Relay	81
6.5.2	Upper Bound for Throughput in the case of Multiple Relays	82
6.5.3	Lower Bound for System Outage	82
6.6	CSI Reduction with Proposed Algorithms	83
6.7	Numerical Results	84
6.7.1	User Generation in the cell	85
6.7.2	Performance Metrics	87
6.7.3	Single RS Case	87
6.7.4	Multiple RS Case	90
6.8	Summary	99

CONTENTS

7	Amplify-and-Forward Cooperative Diversity Schemes for Multi-Carrier Systems	101
7.1	Introduction and Motivation	101
7.2	System Model	103
7.3	Proposed Cooperative Diversity Schemes	105
7.3.1	All-Participate All Subcarrier Scheme (APN) and All-Participate Rate Splitting Scheme (AP-RS)	105
7.3.2	Average Best Relay Selection Scheme (AvgBRS)	106
7.3.3	Per-Subcarrier Best Relay Selection Scheme (NBRS)	107
7.3.4	Random Relay Selection Scheme (RRS)	107
7.4	Preliminary Analysis	108
7.5	Theoretical Analysis of Outage Probability	109
7.5.1	All-Participate All Subcarrier Scheme (APN)	109
7.5.2	All-Participate Rate Splitting Scheme (AP-RS)	111
7.5.3	Average Best Relay Selection Scheme (AvgBRS)	112
7.5.4	Per-Subcarrier Best Relay Selection Scheme (NBRS)	113
7.5.5	Random Relay Selection Scheme (RRS)	114
7.6	Implementation Issues	115
7.7	Numerical Results	115
7.7.1	Simulation results and comparison with the theoretical analysis	115
7.7.2	Simulation results assuming a practical OFDM channel model	118
7.8	Summary	121
8	Conclusions and Future Work	123
8.1	Resource allocation and Scheduling in Cellular MC system	123
8.1.1	Per-Slot Proportional Fair Scheduling (PFS) in a MC System	124
8.1.2	Multi-Slot PFS in a MC System	124
8.1.3	Adaptive Channel State Information (CSI) provision in a MC System	125
8.2	Resource allocation and Scheduling in Relay-based MC system	126
8.3	Cooperative Diversity Schemes for a relay-aided MC System	128
	Appendices	128
A	List of Publications produced during the Ph.D. Studies	129
B	Proof of Theorem 1	133

CONTENTS

Bibliography	135
Curriculum Vitae	143

CONTENTS

Acronyms

2G	<i>2nd</i> Generation
3G	<i>3rd</i> Generation
4G	<i>4th</i> Generation
AF	Amplify-and-Forward
AMC	Adaptive Modulation and Coding
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BS	Base Station
BWA	Broadband Wireless Access
CDF	Cumulative Density Function
CSI	Channel State Information
DF	Decode-and-Forward
DL	Downlink
FDD	Frequency Division Duplex
GPF	General Proportional Fairness
HSDPA	High Speed Downlink Packet Access
LOS	Line Of Sight
MAC	Medium Access Control
MC	Multi-Carrier

ACRONYMS

MIMO	Multiple Input Multiple Output
MRC	Maximal Ratio Combining
MS	Mobile Station
NLOS	Non Line Of Sight
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PF	Proportional Fairness
PFS	Proportional Fair Scheduling
PHY	Physical Layer
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RS	Relay Station
SC	Single Carrier
SINR	Signal-to-Interference plus Noise Ratio
SNR	Signal-to-Noise Ratio
TDD	Time Division Duplex
UL	Uplink
UMTS	Universal Mobile Telecommunications System
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network

Chapter 1

Introduction

1.1 Motivation

The ongoing research efforts in wireless communications are converging towards definition and specification of the 4th Generation (4G) wireless system. The 4G system is expected to be deployed in a wide range of environments and over heterogeneous terminals [1], the ultimate goal being the ubiquitous access to high data rate services. This requires appropriate definition of each of the techniques that will be embedded as building blocks in the future 4G system. The fundamental of a wireless system lies in the usage of the air interface which is regulated by the Physical Layer (PHY). Among the various generic PHY techniques considered for the 4G system, a prominent role is given to the Multi-Carrier (MC) transmission, such as Orthogonal Frequency Division Multiplexing (OFDM) [2]. OFDM is inherently robust to frequency-selective fading and possesses ability to achieve high spectral efficiency [3]. OFDM has been proposed as a physical layer solution for several emerging standards: novel air interface for the evolution of High Speed Downlink Packet Access (HSDPA) in the Universal Mobile Telecommunications System (UMTS) standard [4], the IEEE 802.11a Wireless Local Area Network (WLAN) standard, IEEE 802.16 for high data rate fixed wireless connectivity in larger areas [5] and its mobile version IEEE 802.16e [6], multiband OFDM in the IEEE 802.15.3a for high data rate Wireless Personal Area Network (WPAN), and also in the upcoming IEEE 802.20 standard for achieving data rate of greater than 1 Mbps at 250 km/h, etc.

Besides the PHY technique, the overall architecture supporting the wireless system is an important factor which impacts the performance. Providing high data rates over large areas such as in the traditional cellular system is a difficult problem due to the fundamental trade-off between data rate and cell coverage, for a fixed power and bandwidth. This problem becomes even more severe in the 4G system

where it is envisaged to bring the carrier frequency in the 5 GHz region, decreasing the signal wavelength and thus the cell coverage. The straightforward solution for providing high data rate coverage is to cover the whole area by a higher number of cells. However the deployment of a large number of Base Station (BS) is very costly. Another solution to this problem is to make use of relays, where pre-installed devices or Mobile Station (MS) of some users relay a signal to the users experiencing difficulties. This is an attractive solution which offers the possibility to increase data rate and coverage while minimizing the deployment costs, as mobile relays are anyway present and because a fixed Relay Station (RS) is a low-cost device which can be deployed easily [7].

However, the fundamental building blocks that are the PHY technique and the system architecture do not offer ubiquitous high data rate coverage, how advanced each may be. The perceptual system/user performance deeply relies on an intelligent distribution of the radio resources, a task left for definition to the individual operators and thus having the benefit of a certain range of freedom [8]. The main difficulty in performing a proper allocation arises from the variable quality of a wireless link. Traditionally, this inconsistency was regarded as the main drawback of a wireless system, however this view has shifted with the emerging concept of *multi-user diversity*, which takes advantage of the channel variations instead of compensating for it. This perspective opened the path towards the definition of new types of algorithms. A large number of works has tackled on the resource allocation problem for an OFDM system. There are still many open issues concerning this problem, especially when the MC and relayed transmissions are considered together. In order to take full advantage of the given radio resource, allocation and scheduling algorithms should be designed while embracing the various factors which come into play at different levels, such as the user channel variations in time and frequency and over the spatial links offered by the relays at the PHY layer, the incoming packets from the Medium Access Control (MAC) layer, the required Quality of Service (QoS) for each user, etc.

1.2 The Problems Defined

In this thesis, the problem of radio resource allocation and scheduling is analyzed for the next generation wireless system, with a focus on MC and/or relayed transmission in the Downlink (DL) direction, e. g., from the BS to the MS. Several algorithms are designed which attempt to make the best use of the radio resource offered by the system. Usually, a scheduling problem is expressed as an optimization problem which aims at maximizing (or minimizing) one performance metric, given some constraints. Typically, this is a combinatorial problem for which the obtention of the solution is limited by a prohibitive computational burden. This

creates the need for simpler but effective algorithms, should they be intended for practical use as in the targeted 4G system. This is the spirit in which the algorithms are designed in this thesis. Particular attention has been given to the following aspects for algorithm design:

- **Low computational complexity:** as the wireless channel rapidly changes quality, the algorithm should be able to run at very short time periods, which is the reason why the optimal solutions requiring high complexity are hardly practical. Therefore, there is a need for suboptimal algorithms achieving substantially lower complexities compared to the optimal solution.
- **Combination of different performance metrics:** usually, the scheduling algorithms strive to optimize a single performance metric such as system throughput or user fairness. However, improving several performance metrics at the same time turns out to be more beneficial in practice. The problem is that those performance metrics are in general competing ones. For example, maximizing user fairness requires to allocate resources to a user experiencing poor channel conditions, which drops the system throughput. The aim of the proposed algorithms is to provide a good trade-off between given performance measures such that the individual user requirements or fairness are respected as much as possible while ensuring a high performance at the system level. In particular, the main performance measures are the cell throughput, General Proportional Fairness (GPF) metric for fairness, and the system outage probability.
- **Limited CSI feedback:** scheduling over a MC PHY offers the possibility of achieving high system performance since the radio resource is available in both time and frequency domains. However this comes at the cost of increased signalling as the knowledge of CSI is required at the BS scheduler, in order to take full advantage of the small-scale variations of the users' wireless links. Optimally, the user CSI of all the subcarriers should be sent back to the BS via the Uplink (UL) control channel, in each scheduling time frame. The amount of this information becomes too high, especially as the number of users and subcarriers grow. There is thus a trade-off to be realized between the minimization of the CSI and the achieved performance of the scheduler. The required amount of CSI is an important parameter which should be taken into account in the algorithm design.
- **Benchmarking for assessing the proposed algorithms:** The proposed algorithms are validated by comparison with the optimal solution for each of the target performance metrics. Since the optimal solution is often too complex to compute for large numbers of users/subcarriers, algorithms which

solution approaches the optimal one with reasonable complexity are required.

The allocation problem can be further specified considering the two main system assumptions described below.

1.2.1 Radio Resource Allocation and Scheduling in a MC Cellular System

The problem of resource allocation over an OFDM system has been extensively studied in the literature. In most of the works, the scheduling problem is object of an optimization problem and suboptimal algorithms are proposed. One of the most promising algorithms is the Proportional Fair Scheduling (PFS) algorithm, which aims at optimizing the Proportional Fairness (PF) metric, a widely used objective function, but solved with suboptimal algorithms. The generic idea of the PFS algorithm is to allocate a user with the best instantaneous rate normalized by its average throughput calculated as the throughput allocated to that user over a past window of length T [9]. That is, users with good channels but who did not receive enough data in the past are prioritized. After every slot allocation, the average user rates are updated to account for the newly allocated slot. The mechanism of the sliding window is illustrated in Fig. 1.1, for a system with 4 subcarriers and 4 users. When T is very large, a user with a statistically stronger channel will experience a higher average throughput. With the normalization of the instantaneous rate by the past average throughput, a user is scheduled only when his instantaneous channel quality is high *relative* to his average channel condition over T , e. g., when its channel is near its own peaks [9], thereby ensuring fairness among users. When T is very large and all the users have the same channel statistics, the average throughput of all users becomes the same and PFS is equivalent to the *Maximum CSI* algorithm where the user with the best instantaneous channel is allocated.

However, among the practical algorithms, the suboptimal algorithms based on PFS already proposed have a very poor throughput performance. The achieved throughput should be improved without sacrificing user fairness. Moreover, almost all of these works adopt a very basic system constraint where the scheduling is performed on a slot-by-slot basis. However, in a practical system such as the ones defined in the standards, each scheduling time frame is composed of several time slots, which requires the scheduling over multiple slots in advance. This offers the possibility to exploit the additional degree of freedom in time, by allocating several users over one frequency or subcarrier. Exploiting this new dimension remains an open issue. While scheduling performance is expected to increase, it is

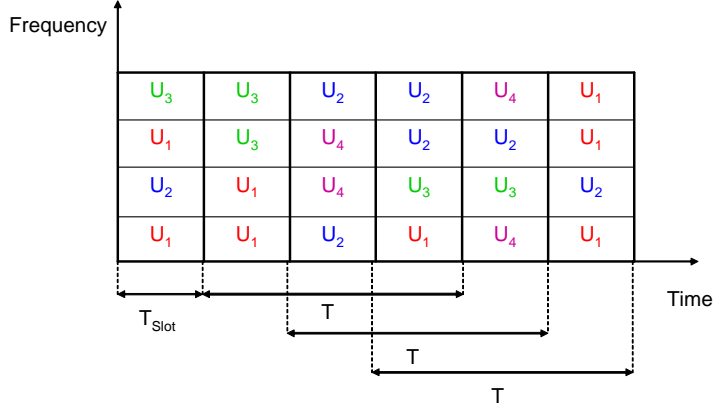


Figure 1.1: Proportional Fair Scheduling: averaging the user rates over the sliding window T

not clear how users should be allocated for realizing the throughput/PF trade-off as well as how much this gain would be.

1.2.2 Radio Resource Allocation and Scheduling in a MC Relay System

Literature on the allocation of radio resource in a MC relay system is scarce. The problem becomes very complex as the allocation options are multiplied with the introduction of relays. Existing algorithms for MC system are not directly applicable in a relayed system and new algorithms are needed, since a relay is neither source nor sink for the communication traffic, while the radio resource allocation for the relay can be done either by the BS or by the RS itself. The optimal allocation supposes that all the decisions are taken at the BS, with a complete freedom in the distribution of the time/frequency resources among the direct or relayed links. However complete freedom is not achievable due to practical hardware limitations which impose some constraints on the radio frame structure. The basic frame structures that can be found in the standards lack in flexibility. To guarantee an efficient resource allocation, the frame itself should be defined along with the allocation algorithm. Moreover, the problem of CSI feedback is amplified as the amount of information is proportional to the number of relays in the system.

It becomes crucial, especially in this case, to design intelligent algorithms which provide a good performance with a minimal amount of CSI feedback. How the radio frame should be divided between the direct (BS to MS), feeder (BS to RS) or relayed (RS to MS) links; how users should be allocated to the direct or relayed links; which entity should perform the scheduling, which users should feedback CSI to which entity, all these interrogations remain open issues.

There are even less studies concerning allocation algorithms in a MC relay system using cooperative diversity, a technique where different relays cooperate to deliver one signal to the destination, thereby increasing transmission reliability. While many works can be found for a Single Carrier (SC) system, the benefit of this technique in a MC system is not clear. This issue first requires thorough analysis from the theoretical point of view with simple assumptions. Allocation schemes which take advantage of the MC aspect should be defined and evaluated, both theoretically and under realistic system assumptions.

1.3 Thesis Contributions

The contributions of the thesis mainly cover two different perspectives: the design of practical algorithms and CSI feedback, and bounding approaches of the optimal solutions for validation. The contributions are organized in a progressive way, following the different system assumptions that are considered:

1. **Algorithms for MC system with per-slot scheduling:** low complexity algorithms achieving a good trade-off between system throughput and PF are designed for the basic case where scheduling is performed for each time slot. To validate the proposed algorithms, a bounding method for approaching the optimal PF solution is developed. With this method, tight bounds to the optimal solution can be derived with reasonable complexity, setting the ground for the validation of the proposed suboptimal algorithms.
2. **Algorithms for MC system with multi-slot scheduling:** having multiple scheduling time slots per frame offers the possibility to exploit the additional degree of freedom in time, by allocating several users over one frequency or subcarrier, a mechanism which will be referred to as *User Multiplexing*. However, the amount of information required for signalling the decision about the user mapping on the DL frame increases with *User Multiplexing*. The proposed algorithms are designed in such a way that minimizes this signalling while optimizing both throughput and PF. It is also shown that User Multiplexing reduces the maximum starvation time, defined as the time where a user does not receive any data. The optimal PF solution in a

MC system is obtained via the same bounding method but this time a tighter lower bound is found by exploiting the allocation over the multiple slots.

3. **CSI encoding scheme:** theoretically, each user sends back his experienced channel quality per subcarrier, per time unit. This is the safest but most inefficient way for feedback, as only the CSI for users who are scheduled is useful, while the CSI of the users having very poor channel condition is useless, since it is unlikely that they will be scheduled. Based on this observation, a flexible CSI encoding scheme is designed which can adapt to the varying needs of the scheduler depending on the number of users in the cell or the type of the scheduler itself.
4. **Algorithms for MC relayed system:** as mentioned earlier, MC transmission and relays are expected to become an integral part of the next generation wireless system. In a MC relayed system, there are so many degrees of freedom that the problem of resource allocation and scheduling becomes prohibitively complex. Algorithms are first designed for a single relay system and second, for a system with multiple relays within a cell. These algorithms take into account the practical assumptions imposed on the frame design and provide a guaranteed per user throughput while achieving a high system throughput at the same time. The optimal solution in terms of either system throughput or outage is also developed, when no practical limitations from hardware or frame architecture are considered.
5. **Generic cooperative diversity schemes in relayed system:** several cooperative diversity schemes are designed for a multiple relay, MC system assuming single user. The outage probabilities of these schemes are mathematically intractable, therefore a bounding approach is taken by means of approximations. This bounding method resulting from the theoretical analysis is validated by comparison with the simulation results, assuming the simplest MC system. The proposed schemes are then applied to an OFDM system. The results confirm the conclusions given by the theoretical analysis, and the best scheme for practical application is recommended.

1.4 Thesis Outline

After the introduction, a general background is given in Chapter 2 which explains the prerequisite notions for the rest of the thesis. The thesis is divided into two parts, each part being in turn divided into several chapters.

Chapter 3 introduces the bounding method for approaching optimal PF in a MC system. One algorithm is derived by lower bounding but still having a high

complexity. Two suboptimal algorithms which enhance throughput and PF are proposed.

Chapter 4 introduces the practical system assumption where each scheduling frame contains several slots. While the upper bound solution remains the same as in the previous chapter, a tighter lower bound is derived, exploiting the new dimension provided by the multiple slots. While the lower bound solution is applicable, its complexity is still high as well as the amount of control information needed for signalling the user mapping in each of the slots. Two suboptimal algorithms are designed which exploit the *User Multiplexing* while minimizing the signalling overhead. A very good trade-off in terms of throughput and PF is achieved, and the maximum starvation time experienced by the users is considerably reduced compared to existing algorithms.

Chapter 5 deals with the problem of CSI feedback reduction. Although this problem is treated as stand-alone, the designed schemes are adaptive in the sense that they can be used by various types of schedulers. That is, the amount and nature of the useful CSI is discussed depending on the number of users and the type of scheduler. The proposed adaptive CSI feedback encoding scheme adapts itself to the varying system requirements, which vary among schedulers, and provides a good trade-off between system performance and feedback reduction.

Chapter 6 introduces the problem of resource allocation and scheduling in a MC system with one or multiple relays within a cell. The increased freedom due to the adjunction of relays that multiply the scheduling options are analyzed. Sub-optimal algorithms for different frame structures and different levels of complexity are designed. Optimal solutions for throughput and outage are also provided, when no system constraints are imposed.

Chapter 7 introduces cooperative diversity in a MC system with relays. Generic allocation schemes making use of cooperative diversity are proposed and theoretically analyzed. The mathematical derivations approximate the actual outage probability of each scheme. The schemes are evaluated assuming a practical OFDM system.

The last chapter provides overall conclusions and directions for future work.

Chapter 2

Background

2.1 Introduction

This part provides a brief overview of the main notions used in this thesis. After introducing the two system architectures considered in the thesis, backgrounds on multi-user diversity, MC transmission, PFS, CSI reporting and relays are given. The system assumptions taken throughout the thesis are then specified. Finally, related works on resource allocation in MC system, CSI reducing schemes, relaying and cooperative diversity schemes are listed.

2.2 System Architecture

The whole thesis is focused on the design of resource allocation and scheduling algorithms for the DL transmissions in a single-cell environment, where the system is based on MC or Orthogonal Frequency Division Multiple Access (OFDMA) technology. Fig. 2.1 shows the system considered in chapters 3, 4 and 5, a single-cell with one BS and multiple users. Each user tracks its channel variations and sends back its measured CSI to the BS. In the most general case, the CSI consists of the Signal-to-Noise Ratio (SNR) values for every subcarrier and scheduling frame, but this definition may be changed. For example, in the case of per subchannel reporting, the CSI can be defined as the average SNR values of the subcarriers belonging to the subchannel. The scheduler located at the BS performs Adaptive Modulation and Coding (AMC), e. g., the modulation rates and coding applied to each subcarrier/subchannel is adapted based on the CSI reports.

The second type of architecture, considered in chapters 6 and 7, is depicted in Fig. 6.3, which is a single cell with multiple relays. In Chapter 6, algorithms are designed assuming multiple users in the cell, whereas Chapter 7 deals with a multiple relay environment with a single user.

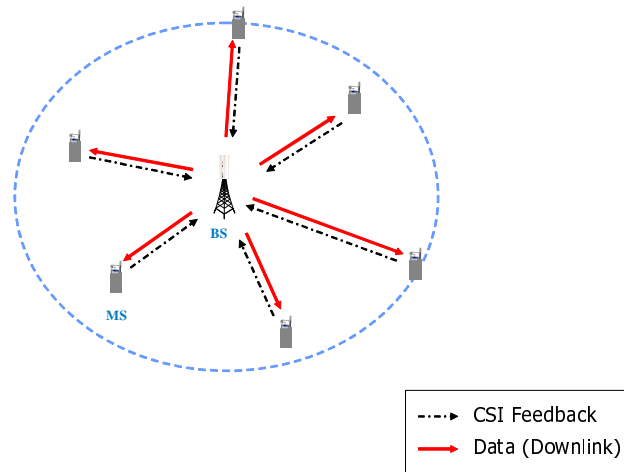


Figure 2.1: First architecture: Cellular System with a Single Cell

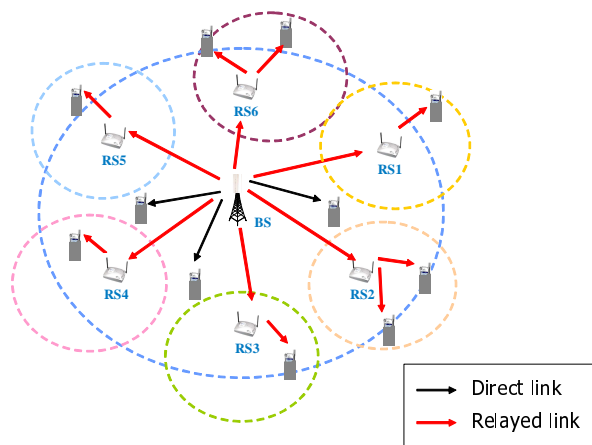


Figure 2.2: Second architecture: Cellular System with Multiple Relays

2.3 Multi–User Diversity

Multi–user diversity is an inherent form of diversity in a multi–user wireless network, provided by the independent fading channels across the different users. Unlike the traditional forms of diversity in time, frequency and space where the channel variations are averaged out on a per user basis, multi–user diversity considers each user channel as a source of randomization and takes advantage of the channel variations observed in the whole system. This idea was first motivated by Knopp and Humblet in [10], where UL transmissions in a single–channel system were considered. Under the assumption that the channel states were measured at the receiver and fed back to the transmitters, it was shown that the information theoretic capacity could be maximized by scheduling at any one time only the user with the best channel gain. The gain from multi–user diversity results from the fact that at any one time, there is likely to be a user whose channel is near its peak. Long–term total throughput can be maximized by always serving the user with the strongest channel [11]. Similar results were obtained for the DL transmissions, from the BS to the MS [12], under the assumption that both transmitter and receivers have perfect channel knowledge. However, maximizing the total throughput only leads to an unfair system as users with constantly poorer channel conditions may never be scheduled. That is, not all users may be fading fast enough to ensure that they will be scheduled in a fair manner, especially in a slow–fading environment. This can result in unacceptable delays for some users to deliver their packets. Thus, an appropriate scheduling is needed in order to exploit multi–user diversity while providing fairness among users, or satisfying some QoS criterion. The challenge is to exploit multi–user diversity while sharing the radio resource fairly and timely to the users with different channel qualities.

2.4 Multi–Carrier System

MC transmission such as OFDM technology is a very promising candidate for the physical layer of the upcoming 4G wireless system due to its high achievable spectral efficiency by bit loading and inherent robustness against inter–symbol interference caused by multi–path fading channels. In a MC system such as OFDM, the frequency–selective wideband channel is divided into independently fading subcarriers, which can be assigned to different users. The variation of the channel in frequency provides an additional source of randomization compared to the time–varying flat fading case, and thus an additional source of multi–user diversity gain. At any moment, over each subcarrier or group of subcarriers, there will be a user with the best channel gain, and as each user experiences different fading channels, the frequency region where a user has the best channel gain will dif-

fer according to the user. Therefore, several users can be scheduled at any one time, each user being assigned the frequency region where his channel gain is the best. This is how multi-user diversity gain is obtained with OFDM. In [11], it is shown that the overall cell throughput gain is maximized by scheduling the user with the best channel gain over each subcarrier, at any one time. However, as in the time-varying flat fading case, this scheme creates unfairness problems among users.

Recently, OFDM has been designated as the PHY layer technology of a number of standards, such as the IEEE 802.16 standard, one of the most promising technologies for providing Broadband Wireless Access (BWA), enabling high data rate services over large areas. The Worldwide Interoperability for Microwave Access (WiMAX) Forum streamlines the implementation of IEEE 802.16 standards [13]. 2 aspects of the OFDM systems have been specified: first, the OFDM option where a single user is allocated the whole bandwidth, and second, the OFDMA option where the subcarriers are grouped into subchannels which can be given to different users. Within the OFDMA option, subchannels can be defined in roughly 2 different ways, the distributed subcarrier configuration, where the subcarriers of one subchannel are spread over the whole spectrum, averaging out the frequency selective fading, and the adjacent subcarrier configuration, employed by the AMC mode, where adjacent subcarriers form a subchannel [14]. With the distributed mode, all users experience similar channel qualities in all subchannels. Thus, resource allocation and scheduling can operate blindly to instantaneous user link qualities. While allocation algorithms are drastically simplified in that case, previous studies have shown very low efficiency of these static allocation schemes [15], since they prevent from any multi-user diversity benefit in the frequency domain. In this thesis, a subchannel is defined as a group of adjacent subcarriers, similarly to the AMC mode of operation in IEEE 802.16, for which radio resource allocation and scheduling issues become much more complex and challenging.

2.5 Proportional Fair Scheduling

One scheduler achieving a trade-off between the multi-user diversity gain and user fairness is the PFS. In the single carrier case, PFS was implemented for Qualcomm's HDR system, also known as $1\times\text{EV-DO}$ [16], where transmission is made when the user's channel condition is favorable, relatively to the user's past condition over a predefined time window. By denoting $r_k(t)$ the current achievable rate of user k in time slot t , and $R'_k(t)$ his past average rate until $t - 1$ allocated

over the time window T , the user with the best *user PFS metric* $\rho_k(t)$, defined as

$$\rho_k(t) = \frac{r_k(t)}{R'_k(t)} \quad (2.1)$$

is allocated in the current slot t . The average rate of each user $R_k(t)$ is calculated as

$$R_k(t) = \frac{(T-1)R'_k(t) + c_k r_k(t)}{T}, \quad (2.2)$$

where c_k is equal to 1 if user k was scheduled in slot t and 0 otherwise. Then, the past average rates $R'_k(t)$ are updated for the next frame,

$$R'_k(t+1) = R_k(t). \quad (2.3)$$

The trade-off is achieved by prioritizing the user with the best achievable rate $r_k(t)$ but with the smallest past average rate $R'_k(t)$. This avoids scheduling the best user at any time since the resource allocated in the past is also taken into account.

PFS defined for the SC system was first adapted to the MC case in a straightforward but suboptimal way, in a number of works such as [17] and [18]. Simply, the SC-PFS is applied in each subcarrier, e. g., each subcarrier n is allocated to the user $k^*(n)$ satisfying

$$k^*(n) = \arg \max_k \rho_{k,n}(t) \quad (2.4)$$

where the user PFS metric $\rho_{k,n}$ is

$$\rho_{k,n}(t) = \frac{r_{k,n}(t)}{R'_k(t)}. \quad (2.5)$$

At time t , $r_{k,n}(t)$ is the achievable rate of user k in subcarrier n and $R'_k(t)$ is the past average rate of user k , summed over all subcarriers and averaged over a window of length T . After all subcarriers are allocated, the average user rates are updated as

$$R_k(t) = (1 - \frac{1}{T})R'_k(t) + \frac{1}{T} \sum_{n=1}^N c_{k,n} r_{k,n}(t), \quad (2.6)$$

where N is the total number of subcarriers. $c_{k,n} = 1$ if subcarrier n is allocated to user k , otherwise $c_{k,n} = 0$. In the following time $t+1$, the past average rate is updated as: $R'_k(t+1) = R_k(t)$. Throughout this thesis, this basic algorithm will serve as a performance benchmark and will be referred as the *Conventional MC-PFS*. This is a suboptimal solution as each subcarrier allocation is made independently from each other.

The condition for achieving optimal MC-PFS is derived in [19], which stipulates that a scheduler P is proportionally fair in a MC system if and only if it maximizes the sum of logarithmic average user rates for scheduler σ , $R_k^{(\sigma)}(t)$, defined in (2.6)

$$P(t) = \arg \max_{\sigma} \sum_{k=1}^K \log R_k^{(\sigma)}(t). \quad (2.7)$$

A scheduler σ is defined as an entity that selects a set of users for a given time and K is the total number of users. We here define as the PF metric of a scheduler σ the GPF parameter, $\Gamma_{(\sigma)}$, as

$$\Gamma_{(\sigma)}(t) = \sum_{k=1}^K \log R_k^{(\sigma)}(t). \quad (2.8)$$

A closed-form expression satisfying this optimality condition is derived in [19]. It is proved that, a scheduler P is PF in a MC system if and only if it satisfies

$$P(t) = \arg \max_{\sigma} \prod_{k \in U_{\sigma}} \left[1 + \frac{\sum_{n=1}^N c_{k,n} r_{k,n}(t)}{(T-1) R_k^{(\sigma)}(t)} \right], \quad (2.9)$$

where U_{σ} is the set of selected users by scheduler σ .

2.6 Channel State Information

To exploit multi-user diversity, the scheduler at the BS needs to know the channel quality of each user. Since the quality of the wireless links are highly variable over small fractions of time, depending on the user speeds and the wireless environment, providing the information about the user channel qualities or CSI becomes a big issue. In general, there are two types of system: Time Division Duplex (TDD) and Frequency Division Duplex (FDD). In a TDD system, the UL and DL transmissions are separated in time but share the same frequency band, so that UL and DL channels can be assumed the same by channel reciprocity, provided that the channel varies slowly enough to have a constant quality over a frame. However, not all users may communicate in the UL direction in all time frames, so their CSI is unknown. In a FDD system, channel reciprocity does not hold as UL and DL transmissions occur in different frequency bands. Thus, in any case, each user needs to feedback his CSI to the BS, which increases the amount of signalling in the UL direction. This becomes especially problematic in a MC system as the CSI for each subcarrier is basically required for each time frame. There is a trade-off to be made between the amount of signalling and the resulting scheduling quality based on the CSI feedback.

2.7 Relay System: Conventional Relaying and Cooperative Diversity

In order to cope with the high data rates envisioned for the 4G system, a fundamental change in the traditional cellular architecture is required, due to two basic reasons. First, the transmission rates envisioned for 4G systems are two orders of magnitude higher than those of 3rd Generation (3G) systems, creating serious power concerns since for a given transmit power level, the symbol (and thus bit) energy decreases linearly with the increasing transmission rate [7]. Second, the 4G spectrum, which is likely to be allocated around the 5 GHz band, results in a radio propagation more vulnerable to Non Line Of Sight (NLOS) conditions. The integration of *relaying* or *multihop* capability into conventional wireless networks seems the most promising architectural upgrade for enabling the ambitious throughput and coverage enhancements, due to the following reasons.

1. Fixed relays are low cost devices: they are only expected to cover regions of diameter 200–500 m, so that transmit power requirements for such a relay are significantly reduced compared to those for a BS. Moreover, the mast on which the relay is placed does not need to be as high as for a BS, reducing operating expenses such as tower leasing and maintenance costs for the service provider [7]. This means easy deployment of a relayed network.
2. Relays do not have a wired connection to the backhaul, thus eliminating the costs of the interface between the BS and wired backhaul network. Instead, they communicate wirelessly with the BS and MS.
3. With the deployment of relays, there is a higher probability that a MS (for example, near the cell edge) is located closer to the RS than the BS, providing links with higher Signal-to-Interference plus Noise Ratio (SINR) over which higher order modulation and code rates can be supported. Consequently, higher per user throughput can be expected as users within the coverage area of the RS receive data, thereby potentially solving the coverage problem for high data rates in large cells. Alternatively, the increased SINR may be used to improve the link reliability [20]. If the loss due to the additional radio resources needed to support the relayed links can be overcome by the throughput gain from the relays, the overall cell throughput can be increased.
4. Relays can be deployed in strategic places so as to provide coverage to users in a coverage hole, e. g., in the shadow of a building or in isolated areas out of reach of any BS. Another example is extending coverage to MS riding on mobile vehicles (e. g., buses and trains), by mounting a RS on the vehicle.

In the existing 2nd Generation (2G) and 3G cellular systems, only analog repeaters with limited capabilities have been used. Most existing and standardized systems were designed for bidirectional communication between a central BS and MS directly linked to them [7]. The deployment of more advanced relaying concepts was hindered by the need for additional radio resources, required to accommodate the additional communication traffic via the relay. Presently, standardization of relay strategies and technologies is considered by the IEEE 802.16's Relay Task Group to enhance the performance of 802.16-based products [21]. Fixed or mobile RS will be owned, installed, and operated by an 802.16 service provider. While detailed discussions are still undergoing over the draft 802.16j, the primary target of this addendum is to provide coverage enhancement (both range extension and mitigation of severe shadowing including coverage hole) and throughput improvement for users at the fringes of an 802.16 cell, as previous 802.16 contributions have shown that such users may have to communicate at reduced information rates due to lower received signal strength. In that sense, throughput enhancement is a consequence of the coverage enhancement, whereas system capacity improvement is not regarded as a primary target of the 802.16j project [21].

In general, relaying systems can be classified as either Amplify-and-Forward (AF) or Decode-and-Forward (DF) systems [7]. In AF systems, the relays amplify the received signal without decoding it. In DF systems, the signal is fully decoded and re-encoded prior to retransmission. There are several modes of operation in a relay system, all of which 802.16j is expected to be compatible with: *conventional relaying*, where the MS only receives the signal forwarded by the RS, or using *cooperative diversity*, where the MS can combine several signals coming from different spatial locations. Cooperative diversity is a form of space diversity which arises by exploiting the following two main features of the wireless medium: the broadcast nature and the spatial independence of its channels [7]. When a source node S sends a message, several relays receive and forward its processed version towards the destination node D . The destination node then combines the signals received from the relays, and also the original signal directly received from the source. By performing this combining, the inherent diversity of the relay channel is usefully exploited [7]. Cooperative protocols can be static, as in the *fixed relaying* scheme defined in [22] where the relays always consume half of the resource for forwarding the incoming signal, or adaptive as in the *selection relaying* scheme of [22], where relays only forward the signal when the source/relay channel gain reaches above a predefined threshold. However these two protocols make inefficient use of the degrees of freedom of the channel, because the relays always repeat, regardless of the direct transmission. By allowing limited feedback from the destination terminal as in the *incremental relaying* scheme [22], e. g., a single bit indicating the success or failure of the direct trans-

mission, the spectral efficiency can be drastically improved, since repetition is made only when they believe it to be helpful for the destination node. More efficient protocols may be defined by allowing the feedback of CSI to the source and relay nodes, while the question of CSI availability in relayed systems is an issue of primary importance.

2.8 System Assumptions

The whole thesis is focused on the design of resource allocation and scheduling algorithms for the DL transmissions in a single-cell environment, where the system is based on MC or OFDMA technology. In chapters 3, 4, and 5 a single-cell system with one BS and multiple users is considered. In Chapter 6, relays are introduced and algorithms are designed assuming multiple users in the cell, whereas Chapter 7 deals with a multiple relay environment with a single user.

2.8.1 Assumption on AMC

Users feed back to the BS their CSI, which consists of the SNR or the corresponding AMC level, (equivalent to the achievable rate), for each subcarrier or each subchannel, defined as a group of adjacent subcarriers. Based on the user CSI, the BS allocates the user data to the radio resource according to a scheduling algorithm. AMC is then applied at each subcarrier or subchannel and for each user. After OFDM modulation, the data frame is sent through the DL channel to all the users. Throughout the thesis, we consider a discrete AMC model, shown in Table 2.1. Discrete AMC levels are used, where SNR thresholds for switching between different AMC modes were determined for a target Bit Error Rate (BER) of 10^{-6} for uncoded M-Quadrature Amplitude Modulation (QAM) symbols, with $M \in \{2, 4, 16, 64, 256\}$ [23] [24].

Table 2.1: Discrete AMC Model

Modulation	BPSK	4-QAM	16-QAM	64-QAM	256-QAM
Rate [b/symbol]	1	2	4	6	8
SNR Threshold [dB]	-5	13.6	20.6	26.8	32.9

2.8.2 Assumptions on Interference

The goal of this thesis is to design scheduling algorithms, under different system assumptions, which allocate OFDMA resource within a single cell, based on CSI reporting. Therefore inter-cell interference (i. e., multiple cell environment) is not considered here. However, intra-cell interference between simultaneously transmitting relays in the cell will be accounted for in Chapter 6. Existing techniques for interference avoidance or frequency reuse can be used on top of the proposed resource allocation. Besides, in multiple cell environment, the CSI reported to the BS is no longer the SNR of the user/BS link but becomes the SINR which reflects the degradation due to interference. Hence, the content of the CSI changes but the principle of the presented algorithms remains the same; only their absolute performance may be lower. However, since the reference algorithms—which also rely on CSI reporting—are deteriorated likewise, our results can serve to infer the relative gain.

2.8.3 Assumptions on Power Allocation

The optimal power allocation strategy over varying channels is given by the well-known *Waterfilling* algorithm, where more power is poured into channels with higher gains [10]. However, it was shown in a number of works that optimal power allocation does not yield so much gain while its required complexity is high [11] [25]. On the contrary, adaptive subcarrier allocation appears to provide much higher gains than adaptive power allocation [26]. In this work, the resource allocation problem will be limited to the problem of subchannels and/or time slots allocation. Power is assumed to be equally distributed over the subcarriers.

2.9 Related Work

A large amount of literature can be found concerning resource allocation and scheduling for OFDMA system. There is a special interest in designing resource allocation schemes in OFDMA system as it offers the possibility to exploit multi-user diversity gain on a per-subcarrier basis, but the scheduling problem becomes more challenging due to this additional frequency dimension. In [11], it is shown that the cell throughput is maximized by scheduling each user on subcarriers where he experiences high channel gain. A large number of works can be found where the scheduling problem is formulated in either of the following ways:

- Maximize the overall data rate under a total transmit power constraint
- Minimize the total transmit power for a given rate requirement,

which are in fact dual problems. The optimal subcarrier allocation for this problem is given in [15], where minimum user rate requirements are considered. Since the optimal algorithm has a prohibitively high complexity, suboptimal algorithms are designed in [15], [27], [28] and [29].

PFS, implemented in Qualcomm's HDR system for single-carrier [16], was adapted to the MC case in [17] and the utility-based optimization approach of [18] derives the same solution. This is a suboptimal solution as each subcarrier allocation is made independently from each other. The optimal PF solution is derived in [19], which shows that each subcarrier allocation depends on the other subcarrier allocations in the same symbol. Other suboptimal algorithms are proposed in [30]. However, those works consider only the optimization of the PFS metric which results in a low cell throughput. Moreover, all these works consider that the scheduling is made slot by slot, e. g., that the basic unit for scheduling or time frame is composed of only one OFDM slot (symbol duration). This is not the case in a practical system such as IEEE 802.16 [5], where in order to limit the computational burden, the allocation overhead, as well as the amount of feedback from the terminals, the scheduling time frame is composed of several OFDM slots.

A large amount of research was also devoted to the problem of reducing the CSI feedback. The most straightforward way to limit this signalling is to feedback only the corresponding AMC levels instead of the SINR values, or feedback on a per-subchannel basis rather than per-subcarrier basis, where a subchannel is defined as a group of contiguous subcarriers. [31] and [32] discuss the impact of the granularity of the time/frequency bin for which CSI is reported on the spectral efficiency. In [33], the throughput behaviour of dynamic allocation with the cost of the signalling overhead is studied, and several signalling schemes are compared. It is shown that despite the overhead, dynamic allocation remains always superior to static schemes. [34] and [35] show that not all the feedback is required to keep the multi-user diversity gain in scheduling. Thus, the AMC level of a subchannel should be reported only if it is higher than a predefined threshold, since this subchannel has a higher probability to be scheduled. [34] and [35] showed that this threshold is dependent on the number of users in the system.

The problem of resource allocation and scheduling for relay-aided cellular systems has been a flourishing topic for investigation and has motivated a number of works such as [36] [37] [38]. However, these works do not consider physical layer based on OFDM. There is a need for a new set of algorithms considering the requirements imposed by the combination of a multiple relay and MC system.

Concerning cooperative diversity, [39] [40] showed that the capacity and robustness of a two user system increases when each user acts as relays for one another, which they refer to as user cooperation. In [22], several cooperative diversity protocols for a single-relay system are proposed and their outage performance is theoretically analyzed. In order to exploit this form of space diver-

sity, many studies have been conducted for the design of the physical layer, in particular the distributed space–time codes as in [41] for single–carrier system, or [42] [43] for an OFDM system with a single relay. In [44], a subcarrier re-ordering scheme is proposed for a single relay system using OFDM transmission. In [45], the optimum relay position for minimizing the outage probability is investigated. However, there are still unresolved problems when taking the higher layer point of view, such as the choice of the relays to be allocated, especially in a MC system.

2.10 Summary

General explanations on the key notions used in this thesis were given. By exploiting the multi–user diversity effect which arises in a multi–user environment, system throughput can be increased. Due to the variations in the frequency domain which occur in MC systems, even higher multi–user diversity gain is obtained. To ensure fairness among users, the PFS has been largely used in SC systems, but the optimal PFS for MC systems requires a much higher complexity. Besides, a problematic point in MC scheduling lies in the CSI reporting for which a lot of UL resource may be wasted. Reducing the amount of CSI is a crucial issue. With the introduction of relays, the problem of resource allocation becomes even more complex, as the degree of freedom of the allocation parameters increases, such as the choice between a direct or relayed link, but at the same time new constraints are imposed, such as the need to spend resources to support the relays. Finally, general assumptions taken over the system were presented, and related works were summarized.

Part I

Radio Resource Allocation and Scheduling in Cellular Multi–Carrier System

Chapter 3

Proportional Fairness in a Multi-Carrier System

3.1 Introduction and motivation

In most of the works found in the literature such as [17] and [18], the PFS adopted for MC system is the *Conventional MC-PFS* presented in section 2.5. This is a suboptimal solution as each subcarrier allocation is made independently from each other. The optimal PF solution is derived in [19], which stipulates that a scheduler is proportionally fair in a MC system if and only if it maximizes the sum of logarithmic average user rates as shown in (2.7) in section 2.5. [19] shows that the optimality condition is expressed as in (2.9) where it is visible that each subcarrier can not be treated independently. That is, the rate allocation made over all subcarriers has to be considered jointly for each user, in order to perform the overall optimization. This is a difficult combinatorial problem with no obvious polynomial time algorithm, which becomes prohibitively complex with a growing number of users and/or subcarriers. Instead, we derive an upper bound on PF by relaxing the initial problem into a convex optimization problem. This upper bound serves as a performance benchmark.

Some variations over the *Conventional MC-PFS* are proposed in the suboptimal PFS algorithms of [30]. In the *Conventional MC-PFS*, the user past average rates $R'_k(t)$ defined in section 2.5 remain constant during the allocation of all the subcarriers in one slot, and are updated only before going to the next slot. The second algorithm proposed in [30] introduces a dependency between the subcarriers, by updating $R'_k(t)$ after every subcarrier allocation, instead of waiting until all the subcarriers in a slot are allocated. In clear, subcarriers are ordered from 1 to N , then subcarrier 1 is allocated to the user with the best user PFS metric $\rho_{k,1}$

and the average rates are updated as

$$R_k(t) = (1 - \frac{1}{T})R'_k(t) + \frac{1}{T}c_{k,1}r_{k,1}(t). \quad (3.1)$$

and the average rates $R_k(t)$ become the past average rates $R'_k(t)$ before allocating subcarrier 2. This procedure continues until subcarrier N is allocated, and then repeated for the next slot. This way of updating the average user rates per subcarrier will be referred to as *partial updates*, as opposed to updating all the subcarriers at the same time as in *Conventional MC-PFS*. These partial updates, which enable a more refined reordering of the user priorities, may be useful to approach the optimal solution which requires such dependencies between subcarriers. However, the drawback of this algorithm is the arbitrary ordering of the subcarriers which are considered for allocation. This means that if the subcarriers were ordered differently, a different user allocation may be produced, since the user priority depends on the allocation made in the previous subcarrier. Thus, the subcarrier ordering has an impact on the overall performance. To avoid this, we define in our proposed algorithms a definite priority rule for determining the order of subcarriers to be allocated, which results in a unique user mapping.

In this chapter, we focus on the case where the scheduling is made on a OFDM slot-by-slot basis, as commonly made in the literature, for the single cell DL transmissions with OFDMA, where users feed back to the BS their CSI as a per-subcarrier SNR. Next chapter will consider the case where the scheduling is made on a frame basis, as in the practical systems. The discrete AMC model is then applied in each subcarrier using Table 2.1. Here the allocation is also made on a per-subcarrier basis. The upper bound on the achievable MC-PFS, is first derived, and then the lower bound by rounding off the upper bound solution into discrete, feasible values. This approximation algorithm giving the lower bound solution can be used as an algorithm, but it has still a high complexity. Thus, we present two low-complexity algorithms which improve both throughput and PF compared to existing algorithms.

3.2 Bounds of the Optimal MC-PF Metric in the Single-Slot case

3.3 Upper Bound of the Optimal MC-PF Metric

In this section the upper bound of the optimal MC-PF metric is derived. This problem can be written as maximizing the sum of logarithmic average user rates,

e. g., the GPF metric defined in (2.8), subject to certain constraints

$$\begin{aligned} & \text{minimize} && -\sum_{k=1}^K \log[(a^k)^t x + b^k] \\ & \text{subject to} && (h^n)^t x = 1, \quad n = 1, \dots, N \\ & && x_i \in \{0, 1\}, \quad i = 1, \dots, KN \end{aligned} \quad (3.2)$$

where the variable $x \in \mathbf{R}^{KN}$, the set of real-valued vectors of size $KN \times 1$. $(\cdot)^t$ denotes the transposition operator. The superscript $(\cdot)^k$ denotes that a^k and b^k are determined for user k , and h^n is calculated for subcarrier n . The problem instance is defined by the constants $a^k \in \mathbf{R}^{KN}$, $h^n \in \mathbf{R}^{KN}$ and $b^k \in \mathbf{R}$, the set of real values. a_i^k , the i -th element of vector a^k , h_i^n the i -th element of vector h^n , and b^k are defined as

$$a_i^k = \begin{cases} \frac{r_{k,n}}{T}, & \text{for } i = (k-1) \times N + n; n = 1, \dots, N \\ 0 & \text{otherwise} \end{cases}$$

$$h_i^n = \begin{cases} 1, & \text{for } i = k \times n; k = 1, \dots, K \\ 0 & \text{otherwise} \end{cases}$$

$$b^k = \frac{T-1}{T} \times R'_k.$$

Vector x is the allocation vector where its i -th element expresses the fraction of each subcarrier n allocated to each user k . Thus, comparing the objective function of (3.2) with the average rate expression defined in (2.6) in section 2.5, $(a^k)^t x$ stands for the portion of the allocation made in the current slot, whereas b^k is the portion of the past average rates.

In vector notations, x and a^k are written as

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_{K \times N} \end{pmatrix}, \quad a^k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ r_{k,1}/T \\ \vdots \\ r_{k,N}/T \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

h^n is a block vector composed of K blocks of size N . The vector corresponding

to one block of size N , denoted \bar{h}^n , is

$$\bar{h}^n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the j -th element \bar{h}_j^n is equal to 1 if $j = n$ and zero otherwise, e. g.,

$$\bar{h}_j^n = \delta_{j,n}, \text{ for all } j \in \{1..N\}. \quad (3.3)$$

Since h^n is built by concatenating K times the vector \bar{h}^n , e. g.,

$$h^n = \begin{pmatrix} \bar{h}^{n,1} \\ \bar{h}^{n,2} \\ \vdots \\ \bar{h}^{n,K} \end{pmatrix},$$

its i -th element h_i^n is equal to

$$h_i^n = \delta_{i,k \times n}, \text{ for all } k \in \{1..K\}, \text{ for all } i \in \{1..KN\}. \quad (3.4)$$

Vector h^n is created in order to extract from the allocation vector x , the terms corresponding to the allocation of subcarrier n to all users. Thus, the first constraint expresses that only one user can be allocated to each subcarrier n .

However, (3.2) is a difficult combinatorial problem, but if the constraints $x_i \in \{0, 1\}$ are relaxed to $0 \leq x_i \leq 1, i = 1, \dots, KN$ we get a *convex optimization problem*,

$$\begin{aligned} & \text{minimize} && -\sum_{k=1}^K \log((a^k)^t x + b^k) \\ & \text{subject to} && (h^n)^T x = 1, \quad n = 1, \dots, N \\ & && 0 \leq x_i \leq 1, \quad i = 1, \dots, KN \end{aligned} \quad (3.5)$$

which can be solved efficiently (and globally) using an interior point algorithm [46]. The solution to (3.5) naturally gives an *upper bound* to (3.2) since the problem is solved over a larger domain. Such an approach is a *convex relaxation* and has been very successful in, e. g., combinatorial optimization [47]. The relaxation of the discrete coefficients $x_i \in \{0, 1\}$ to continuous ones $0 \leq x_i \leq 1, i = 1, \dots, KN$ can be interpreted as the allocation of a fraction of the subcarrier corresponding to the value x_i . That is, a subcarrier would be shared among different users. This is not possible to implement since it is assumed that there is only a single OFDM symbol per allocation instance. This optimization is performed every time slot. The average user rates are updated for the next slot according to x_i . The upper bound GPF and the achieved throughput are calculated based on x_i .

3.3.1 Lower Bound of the Optimal PF Metric by Rounding the Solution to a Discrete Feasible Point

The first approximation algorithm is derived by means of rounding, referred as *Rounding* algorithm. The solution to (3.5) gives an upper bound to (3.2) and must be mapped to a discrete point satisfying the constraints in (3.2). This can be interpreted as a *lower bound* on the solution to (3.2), the optimal MC-PF, since the discrete point is in general suboptimal.

In *Rounding* algorithm, the terms of x are mapped to either 1 or 0, by assigning a subcarrier n^* to the user k^* for which x_{kn^*} is maximum among all users k : $x_{k^*n^*} = 1$ and $x_{kn^*} = 0$ for $k \neq k^*$. Thus, the subcarrier is assigned to the user with the highest fraction.

3.4 Proposed Approximation Algorithms

3.4.1 Trade-off between Rate and Fairness: Best-Rate and Best-PFS algorithms

While having a near-optimal PF, *Rounding* algorithm does not achieve the best throughput, since the optimization is made only in terms of the GPF metric. Moreover, *Rounding* algorithm still has a high computational complexity. Two other suboptimal algorithms are proposed, which make a trade-off between rate and PF: the *Best-Rate* algorithm which aims at near-optimal rate while keeping a good PF level, and the *Best-PFS* algorithm, which targets a near-optimal PF, at a small expense of the throughput. These algorithms were designed bearing in mind the following factors: the assumed discrete AMC model, the partial average user updates and the need for a definite rule for subcarrier ordering as discussed above.

- **Discrete AMC model:** since a discrete AMC model is assumed, several users with different SNR values may actually have the same achievable rate. This can happen quite often since there are only 5 discrete AMC levels in which the user SNR values are mapped, by comparison with the SNR thresholds from Table 2.1. Thus, users having the same best achievable rate can be sorted out by their user PFS metric, which depends on their average rates served in the past. The proposed algorithms take advantage of this feature by selecting the user k which is simultaneously the best in terms of rate metric $r_{k,n}$ and user PFS metric $p_{k,n}$, for a subcarrier n .
- **Definite rule for subcarrier ordering:** in the proposed algorithms, the subcarriers are not ordered prior to the allocation process, but ordering is integrated within the algorithm and is performed iteratively. In the first

step, in each subcarrier, all users are sorted according to the two metrics: rate and user PFS metric. Then, in each subcarrier, the best users for rate and PFS metric are compared. If the best users for both metrics coincide, e. g., it happens to be the same user, this user is allocated in the corresponding subcarrier. All such subcarriers are allocated within this first step. Thus, this matching criteria, where the allocated user is the best for both metrics, allows a selection of a group of subcarriers where this phenomenon is observed. These selected subcarriers are all labelled "number 1". After this, the average user rates are partially updated by the mechanism explained below, then this matching process is repeated. The subcarriers allocated in this second step can be labelled as "number 2". This process is repeated until all N subcarriers are allocated. Since the number of allocated subcarriers per step may be larger than one, the total number of steps or iterations L will follow $L \leq N$. However, the algorithms are blocked if in one step, no matching between the best rate user and best PFS user occur in any subcarrier. This is where the two algorithms differ. In the *Best-Rate* algorithm, subcarriers are ordered according to the rate of their best rate user. The best subcarrier, e. g., the one supporting the best rate user compared to all the best rate users in the other subcarriers, is allocated. Again, if there are several such best subcarriers, all these subcarriers are allocated with their best rate user. After this allocation process, the average user rates are partially updated and the matching process can be applied again since the user PFS metrics have changed. The *Best-PFS* algorithm works exactly in the same way, except that the user PFS metrics are used instead of the rates. Details of the steps of both algorithms are provided further.

- **The partial average user updates:** as explained above, subcarriers are not necessarily ordered one by one, but rather in group unit, each group of subcarriers being allocated in each step. Thus, if L denotes the total number of iterations, then $L \leq N$. After each l -th iteration, the average user rates are updated. However, performing partial updates after each iteration and making one global update after all the subcarriers are updated as in *Conventional MC-PFS* do not result in the same average user rates. This is obvious from the fact that the past average rates $R'_k(t)$ are multiplied by the coefficient $1 - 1/T$ while the rates allocated in the current step are multiplied by $1/T$. Therefore, the mechanism of *virtual partial* average user rates updates is introduced, as opposed to the *global* updates after all subcarriers have been allocated. The *global* updates, which are simply the ones defined by (2.6) reflect the rates really perceived by the users, whereas the *virtual partial* updates are created only for facilitating the allocation. They work as follows. Initially is set $c_{k,n} = 0$ for all n, k . At the l -th iteration, for each

subcarrier n^* allocated to each user k^* , set $c_{k^*,n^*} = 1$. After the l -th iteration, the allocated users' average rates are partially updated as

$$R_{k^*}^l(t) = (1 - \frac{1}{T})R_{k^*}^{l-1}(t) + \frac{1}{T} \sum_{n=1}^N c_{k^*,n} r_{k^*,n}(t). \quad (3.6)$$

and

$$R_{k^*}^{l+1}(t) = R_{k^*}^l(t). \quad (3.7)$$

Below is provided the detailed explanation of the proposed *Best-Rate* and *Best-PFS* algorithms, which only differ in step 4). The set of empty subcarriers to be allocated is denoted \mathbf{N}_E . Initially, the cardinality $\text{card}(\mathbf{N}_E) = N$.

1) For $n \in \mathbf{N}_E$, order users by:

- order of decreasing rate metric $r_{k,n}$
- order of decreasing user PFS metric $\rho_{k,n}$

If several users have the same rate¹, put them in the same level of order.

2) For $n \in \mathbf{N}_E$, determine the best PFS user $k_{n,\rho}^*$ and the best rate user $k_{n,r}^*$ as:

$$k_{n,\rho}^* = \arg \max_k \rho_{k,n} \quad (3.8)$$

$$k_{n,r}^* = \arg \max_k r_{k,n} \quad (3.9)$$

In each n , compare $k_{n,\rho}^*$ and $k_{n,r}^*$:

- if $k_{n,\rho}^* \neq k_{n,r}^*$: add n in $\mathbf{N}_{E'}$, the set of empty subcarriers after this matching process, and go to the next subcarrier.
- if $k_{n,\rho}^* = k_{n,r}^*$: allocate n to this user and go to the next subcarrier.

Repeat this procedure for every subcarrier.

3) If there were some matching cases, i.e. if $\text{card}(\mathbf{N}_{E'}) < \text{card}(\mathbf{N}_E)$, partially update the average user rates with (3.6) and go to step 5). If there were no matching at all, i.e. if $\text{card}(\mathbf{N}_{E'}) = \text{card}(\mathbf{N}_E)$, proceed to step 4).

4) The algorithms differ in this step:

4A) *Best-Rate algorithm*: select the best rate user k_r^* among all best rate users $k_{n,r}^*$ and his corresponding subcarrier n^* among $n \in \mathbf{N}_{E'}$:

$$(k_r^*, n^*) = \arg \max_{k_{n,r}^*, n \in \mathbf{N}_{E'}} r_{k,n} \quad (3.10)$$

¹This happens often as the user SNRs are mapped to 5 discrete Adaptive Modulation and Coding (AMC) levels by comparison to the corresponding SNR thresholds

Allocate user k_r^* to subcarrier n^* .

4B) *Best-PFS algorithm*: select the best PFS user k_p^* among all best PFS users $k_{n,p}^*$ for $n \in \mathbf{N}_{E'}$:

$$(k_p^*, n^*) = \arg \max_{k_{n,p}^*, n \in \mathbf{N}_{E'}} \rho_{k,n} \quad (3.11)$$

Allocate user k_p^* to subcarrier n^* .

For 4A) and 4B): partially update the average user rates with (3.6) and $\mathbf{N}_{E'} = \mathbf{N}_{E'} - \{n^*\}$.

5) PFS metrics are reordered. Re-initialize the set of remaining subcarriers: $\mathbf{N}_E = \mathbf{N}_{E'}$ and the set of empty subcarriers after matching: $\mathbf{N}_{E'} = \emptyset$. Repeat the algorithm from step 1 if $\mathbf{N}_E \neq \emptyset$ and end when $\mathbf{N}_E = \emptyset$.

With the *partial* updates of steps 3) and 4), the allocation utilizes sequence of subcarriers that is directed towards more optimal PF, unlike in the previous solutions [30].

3.5 Numerical Results

Throughput and GPF performance of PFS algorithms are highly dependent on the averaging window size T . It is stated in [48] that T is related to the *maximum starvation time*, defined as the maximum amount of time for which an individual user can be starved, i. e., not receive service. Besides, the longer T , the higher the throughput since the algorithm can wait longer for a high channel peak, but the longer the maximum starvation period. In that work, T is chosen to be equal to 1.67 seconds, since it gave the best compromise between throughput and starvation period, according to their simulations. However, minimizing T while achieving the same throughput is very beneficial since it will shorten the maximum starvation and thus improve the overall quality of service. Thus, evaluation is made for a relatively small window of 5 slots.

We consider a single cell of 1000m radius with uniformly distributed users and 2000 sets of channel realizations, where the user locations are regenerated for each set of channel realizations. The path loss model for urban areas from [49] is used, for a carrier frequency of 5 GHz, which gives a mean path loss L in [dB]

$$L = 136.51 + 37.6 \times \log_{10}(R_c), \quad (3.12)$$

where R_c is the cell radius expressed in *km*. For the generation of multipath fading, the model from [50] is used. It is assumed that the multipath fading channel has an exponential delay power profile, with each path gain envelop being Rayleigh distributed. Log-normal shadowing with 0 mean and 12 dB standard deviation is assumed. 16 subcarriers and 5 to 20 users are considered. The algorithms are compared with two existing PFS algorithms, Conventional MC-PFS and PFS per

subcarrier, the best one in [30] where each subcarrier is allocated by independent PFS. Max CSI algorithm is represented, where the user with best SNR is allocated each subcarrier, without any PF consideration. Fig. 3.1 and 3.2 show GPF and cell throughput, respectively. In the throughput calculation, actual symbol error rates depending on the AMC level are included, and is obtained by

$$\tau = \sum_k \sum_n r_{k,n} \times (1 - r_{k,n} \times BER_{k,n}), \quad (3.13)$$

where $BER_{k,n}$ is the bit error rate experienced by user k on subcarrier n . Fig. 3.1 shows that Rounding has a near-optimal PF, as the optimal solution is located between the upper-bound and Rounding. However, Fig. 3.2 shows that the throughput of Rounding is outperformed by Best-PFS and Best-Rate, which approaches Max CSI but with much higher PF. This is because the AMC model allows users with different CSI to achieve identical rate. While Max CSI selects the users with the best SNR, Best-Rate selects among the best rate users, the ones with highest PFS. Moreover, the GPF decrease of Max CSI at $K = 20$ is due to the increased number of unscheduled users for constant $N = 16$: after several frames, their average rates decrease and become less than 1, leading to a negative contribution to GPF (see (2.8)). The decrease of the throughput of Rounding for more than 15 users is due to the fact that, the more users, the more Rounding diverges from the upper bound as the difference between the allocated subcarrier fractions increases.

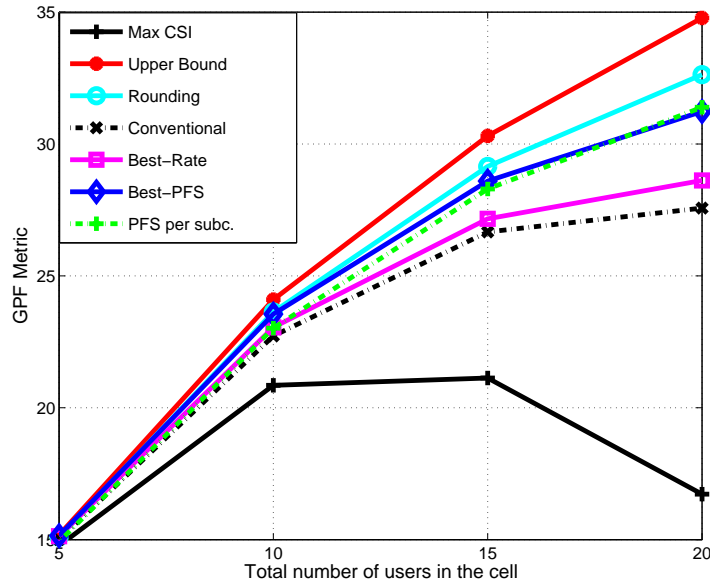


Figure 3.1: Performance in terms of average GPF metric

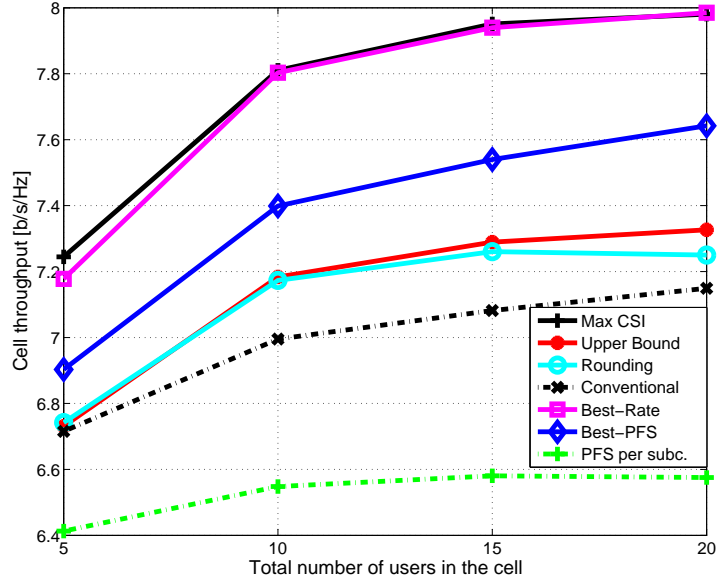


Figure 3.2: Cell throughput in [b/s/Hz]

Compared to Best-Rate, Best-PFS improves PF, at the expense of throughput but still outperforms existing algorithms. Besides, Best-Rate and Best-PFS algorithms delimit a region for each metric. Any point of this region can be achieved by a combined algorithm which selects randomly step 4A) or 4B).

Finally, a complexity analysis shows that Conventional MC-PFS grows as $O(2KN)$, and Rounding algorithm as $O((2KN)^3 + KN)$, since the complexity order involved for the obtention of the PF upper-bound is at least as $O((2KN)^3)$. For α iterations, our algorithms grow as $O(2\alpha KN)$, which equals in the worst case to $O(2KN^2)$ for $\alpha = N$. Both algorithms achieve a better performance than Conventional MC-PFS with a reasonable complexity, much lower than the one required by the optimal solution.

3.6 Summary

In this chapter, we have addressed the problem of PFS in the basic case of slot-by-slot scheduling. The PF upper bound of the optimal MC-PF metric was derived for a fixed PFS window size. Three approximation algorithms have been proposed. The first one achieved a near-optimal PF, but with high complexity. With lower complexity, the two others improved both throughput and PF, compared to Conventional MC-PFS: one achieved a near-optimal throughput and the other increased PF.

Chapter 4

Proportional Fairness in a Multi-Carrier System with Multi-Slot Frames

4.1 Introduction and Motivation

In section 2.5, the optimality condition for PF in MC case derived in [19] was shown. However, it can be seen that realization of optimal PFS in a MC system is a difficult combinatorial problem with no obvious polynomial time algorithm, which becomes prohibitively complex with a growing number of users and/or subcarriers. Moreover, this condition is optimal only if we consider a per-slot scheduling, i. e., if the basic unit for scheduling or time frame is composed of only one OFDM slot (one OFDM symbol)¹. However this is not the case in a practical system such as IEEE 802.16 [5], where in order to limit the computational burden, the allocation overhead, as well as the amount of feedback from the terminals, the time frame is composed of several OFDM slots.

As in the previous chapter, most of the existing allocation algorithms in the literature deal with the per-slot allocation, referred here as the *single-slot case*. Here, we consider a practical time frame containing several OFDM slots and we introduce the idea of scheduling different users on a same subcarrier in a time frame but in different time slots. Indeed, as the basic unit of allocation is defined by a time slot and a subcarrier, different users can be multiplexed on different slots in the same subcarrier, in a time-division manner, referred as *User Multiplexing* or *Sharing*. When there is no *User Multiplexing*, each subcarrier can be allocated to only one user. This is illustrated in Fig. 4.1, where Fig. 4.1 (a) shows a frame

¹In the optimal case, a slot is equal to an OFDM symbol duration, i.e., an OFDM symbol is the smallest allocable unit in time, but in practice, a slot can contain several OFDM symbols.

with User Multiplexing whereas Fig. 4.1 (b) shows an example without User Multiplexing, where the frame contains 4 slots and 2 subcarriers. *User Multiplexing* certainly generalizes the single-slot case with new problems and requirements, as it will become clear throughout this chapter. In particular, we will show that PFS algorithms designed for the single-slot case or the ones proposed in Chapter 3 perform rather poorly in the multi-slot case, when applied slot-by-slot, due to the high overhead induced, and hence new types of algorithms are needed. Thus, the key to our proposal is the optimized user-multiplexing with a low-overhead. First, we show that with multi-slots, the optimal allocation becomes even more complex. Therefore, we determine the upper bound for the optimal PF metric by means of convex optimization and a lower bound is derived by a rounding algorithm. Note that these are the bounds for optimal PF, which do not achieve the optimal throughput, since throughput and PF are two competing measures. As achieving both throughput and PF optimization is very valuable for resource allocation, we propose such user multiplexing algorithms. Their performance is given in terms of throughput and PF, including the overhead due to the descriptonal complexity of such a multiplexing.

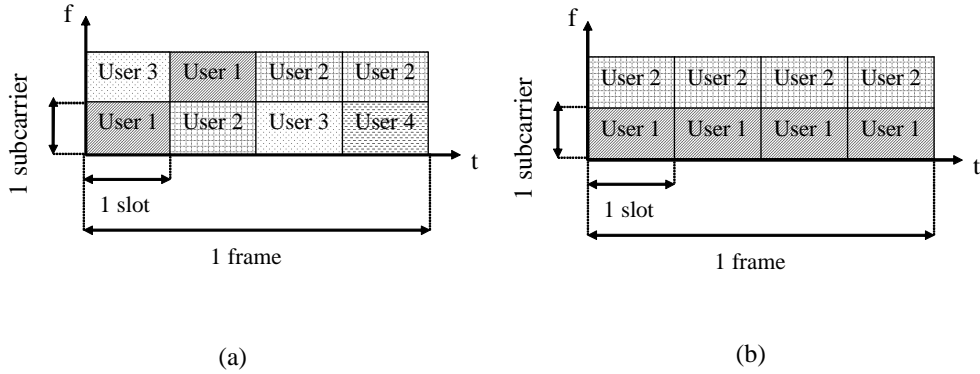


Figure 4.1: Example of allocations in a multi-slot frame: (a) with user multiplexing (b) without user multiplexing

In this chapter, the users in the cell are considered stationary or with low mobility. Thus, the channel condition of each user can be assumed constant for all the slots during a frame. This is a reasonable assumption made in practical systems such as [5]. The discrete AMC model from Table 2.1 is applied.

4.2 Proportional Fair Scheduling for Multi-Carrier System, in the Multi-Slot case

The optimality condition for PF for MC system shown in section 2.5 holds only for the single-slot case, as it creates the constraint of selecting one user per subcarrier, such that the same user is scheduled on all the slots for that subcarrier. But with S slots per frame, up to S different users can use a single subcarrier during a frame, each user occupying one slot. In this case, the optimality condition corresponding to the one in [19] is instead

$$P(t) = \arg \max_{\sigma} \prod_{k \in U_{\sigma}} \left[1 + \frac{\sum_{n=1}^N s_{k,n} r_{k,n}(t)}{(T-S)R_k^{(\sigma)}(t)} \right]. \quad (4.1)$$

In the current frame, $r_{k,n}(t)$ is the achievable rate of user k in subcarrier n and $s_{k,n} \in \{0, 1, 2, \dots, S\}$ is the number of slots allocated to user k on subcarrier n ; N is the total number of subcarriers. T is the PFS window length equal to a multiple of the scheduling time frame, n_f , expressed in terms of number of OFDM slots, $T = n_f \times S$. $R_k^{(\sigma)}(t)$ is the past average rate of user k , summed over all subcarriers and averaged over window T . In the multi-slot case, the average user rate becomes

$$R_k^{(\sigma)}(t) = \frac{(T-S)R_k^{(\sigma)}(t) + \sum_{n=1}^N s_{k,n} r_{k,n}(t)}{T} \quad (4.2)$$

The different formulation of the average rates compared to the single-slot case is due to the weight of the newly allocated rates, $\sum_{n=1}^N s_{k,n} r_{k,n}(t)$ which spans over S slots, so that the contributing fraction from the past average rate $R_k^{(\sigma)}(t)$ becomes $\frac{T-S}{T}$. $R_k^{(\sigma)}(t)$ is updated for the next frame as $R_k^{(\sigma)}(t+1) = R_k^{(\sigma)}(t)$.

By comparing the condition (2.9) given in [19] and (4.1), it becomes clear that the scheduling possibilities are increased compared to the single-slot case, as the degree of freedom in time is introduced. That is, 0 to S slots can be allocated to each user now, whereas it was only possible to allocate 0 or 1 slot per user in the single-slot case. Hence the scheduler for which the optimality condition for single-slot case is maximized will be upper-bounded by the scheduler for which (4.1) is maximized. This statement motivates the concept of User Multiplexing.

For the *Conventional MC-PFS* algorithm applied to the multi-slot case, the user with the best PFS metric occupies all the slots of a subcarrier. Note that the subcarriers are allocated independently from each other, while the optimality conditions [19] and (4.1) imply that a subcarrier allocation is dependent on the allocations made in the other subcarriers.

4.3 Bounds of the Optimal MC-PF Metric in the Multi-Slot case

4.3.1 Upper Bound of the Optimal-PF Metric in the Multi-Slot case

In the previous chapter, the problem for the optimal MC-PFS allocation was formulated for the single-slot case. In the multi-slot case, the constraints on the optimization problem are slightly different and can be written as

$$\begin{aligned} & \text{minimize} && -\sum_{k=1}^K \log[(a^k)^t x + b^k] \\ & \text{subject to} && (h^n)^t x = S, \quad n = 1, \dots, N \\ & && x_i \in \{0, 1, 2, \dots, S\}, \quad i = 1, \dots, KN \end{aligned} \quad (4.3)$$

where the variable $x \in \mathbf{R}^{KN}$ was defined in section 3.2, as well as $h^n \in \mathbf{R}^{KN}$. Vectors a^k and b^k are now defined as

$$a^k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ S \times r_{k,1}/T \\ \vdots \\ S \times r_{k,N}/T \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b^k = \frac{T-S}{T} \times R'_k.$$

The first constraint expresses that the sum of all users' allocation factors x_i , $i = k \times n$ over a subcarrier n is equal to the number of slots per frame, S . The channel state is constant over the S slots in a frame since, as mentioned earlier, the CSI is fed back every frame in a practical system, which implies that $r_{k,n}$ is also constant over the S slots. As in section 3.2, this difficult combinatorial problem is transformed into a convex optimization problem by relaxing the constraints $x_i \in \{0, 1, 2, \dots, S\}$ to $0 \leq x_i \leq S$, $i = 1, \dots, KN$,

$$\begin{aligned} & \text{minimize} && -\sum_{k=1}^K \log[(a^k)^t x + b^k] \\ & \text{subject to} && (h^n)^t x = S, \quad n = 1, \dots, N \\ & && 0 \leq x_i \leq S, \quad i = 1, \dots, KN \end{aligned} \quad (4.4)$$

As in section 3.2, the solution to the relaxation (4.4) gives an *upper bound* to (4.3). While the optimal problems for single-slot and multi-slot cases give completely

different solutions (see the domains of x and the definition of variables), the solution for the relaxation, e. g., the upper bound, will be the same for both cases. This is because the relaxed solution is in both cases interpreted as the allocation of a fraction of the subcarrier corresponding to the value x_i , whether in one slot or S slots. Basically, single-slot is a special case of multi-slot allocation. This optimization is performed every time frame. The average user rates are updated for the next frame according to the allocation obtained by x_i in the current frame. This can be interpreted as the allocation of a fraction of the subcarrier corresponding to the value x_i .

4.3.2 Lower Bound of the Optimal PF Metric by Rounding the Solution to a Discrete Feasible Point in the Multi-Slot case

As explained in the previous chapter, by mapping the solution to the relaxation into discrete points satisfying the constraints in (4.3), we obtain a feasible solution which provides a lower bound on the solution to (4.3). This is referred as *Multi-Slot Rounding*.

In the single-slot case, the rounding approach in section 3.4 was to map the terms of x to either 1 or 0, since each subcarrier can be applied to only one user. In the multi-slot case the rounding approach is generalized. That is, a finer approximation is possible since there are S slots or bins where the real valued x_i can be fractioned, instead of only one. The approach is to round the terms of x_{kn} , the $k \times n$ -th term of solution vector x , to their closest discrete values \bar{x}_{kn} from 0 to S . By doing so, the first constraint may be violated. We want to keep the factors x_{kn} for which $|(x_{kn} - \bar{x}_{kn})| = \Delta_{kn}$ are minimum. There are two cases:

- If the sum of the rounded values is larger than S , $\sum_{k=1}^K \bar{x}_{kn} = S' > S$, it is not feasible since there are only S slots. To make this sum equal to S , we identify the $S' - S$ terms \bar{x}_{kn} with the largest gap with their exact solution x_{kn} , namely the $S' - S$ terms for which Δ_{kn} are highest. These $S' - S$ terms \bar{x}_{kn} are reduced by 1.
- If $\sum_{k=1}^K \bar{x}_{kn} = S' < S$, some slots are empty so the resources are not fully utilized. We increase by 1 the $S' - S$ terms \bar{x}_{kn} with the largest Δ_{kn} , so the S slots are all allocated.

In both cases, the reallocation concerns the terms with the largest gap since for these terms, the rounded values \bar{x}_{kn} are the most distant from the exact solution x_{kn} . Thus, the rounded values closest to the exact solution are kept, while no more and no less than the slots contained in a frame are allocated. Fig. 4.2(a) shows

that the fractions computed for the upper bound do not necessarily correspond to the predefined slots, but in Fig. 4.2(b), the allocation resulting from the Rounding algorithm produces a feasible allocation.

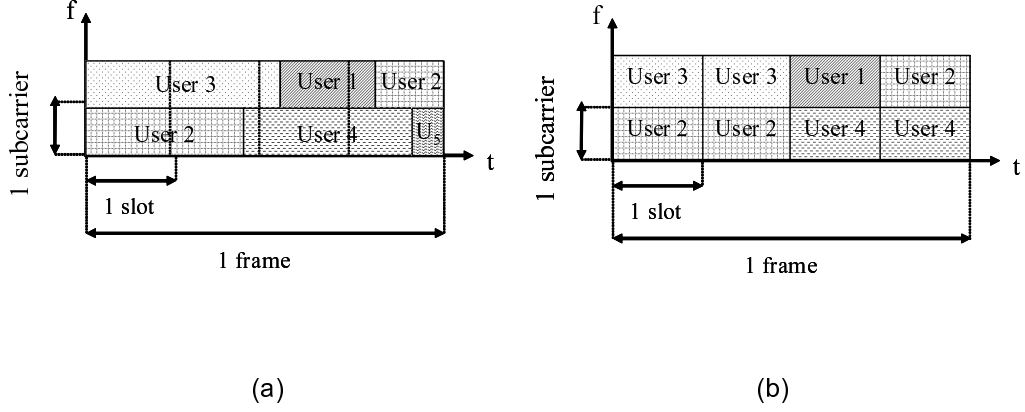


Figure 4.2: Example of allocations in a multi-slot frame: (a) Upper Bound (b) Rounding

4.4 Proposed Algorithms with User Multiplexing

The GPF parameter can be decomposed as follows

$$\Gamma_{(\sigma)}(t) = \sum_{k=1}^K \log\left[\frac{T-S}{T} R_k^{(\sigma)}(t)\right] + \sum_{k=1}^K \log\left[1 + \sum_{n=1}^N \frac{s_{k,n} r_{k,n}}{(T-S) R_k^{(\sigma)}(t)}\right] = A + B_N. \quad (4.5)$$

where

$$B_n = \sum_{k=1}^K \log\left[1 + \sum_{j=1}^n \frac{s_{k,j} r_{k,j}}{(T-S) R_k^{(\sigma)}(t)}\right]. \quad (4.6)$$

A is constant as the past rates R_k' are given during one allocation instance, but B_n varies during subcarrier allocation. For our heuristic algorithms, we introduce the *partial GPF metric* $\Gamma_n = A + B_n$, which represents the GPF metric for $n \leq N$ allocated subcarriers during a frame allocation.

When the allocation of all the subcarriers is finished, the partial GPF metric is equal to the GPF metric. Our algorithms are performed on a per subchannel basis, over the M subchannels. $\bar{r}_{k,m}$ denotes the average rate of the N/M subcarriers in subchannel m .

4.4.1 PF–User Multiplexing Algorithm (PF–Mux)

This algorithm works as follows. In each subchannel m , all the users are ordered according to decreasing user PFS metric $\rho_{k,m}$. The best PFS users are first allocated in each subchannel. At each step $a \in \{1, \dots, \log_2(S)\}$, the number of multiplexed users per subchannel is doubled from 2^{a-1} users to 2^a users if the following two conditions are fulfilled. The first condition is the rate condition

$$\sum_{k=1}^{2^a} \frac{\bar{r}_{k,m}}{2^a} \geq \sum_{k=1}^{2^{a-1}} \frac{\bar{r}_{k,m}}{2^{a-1}}, \quad (4.7)$$

or equivalently

$$\sum_{k=2^{a-1}+1}^{2^a} \bar{r}_{k,m} \geq \sum_{k=1}^{2^{a-1}} \bar{r}_{k,m}, \quad (4.8)$$

where k is ordered by decreasing user PFS metric. This condition ensures that the average rate after sharing 2^a users will be increased compared to the case with 2^{a-1} users. The second condition is the proportional fairness condition

$$B_m^{2^a} \geq B_m^{2^{a-1}}, \quad (4.9)$$

where superscripts 2^a and 2^{a-1} denote that 2^a and 2^{a-1} users are multiplexed, respectively. (4.9) ensures that the partial GPF metric will be increased when allowing 2^a users, since the term A defined in (4.5) is constant during one frame allocation.

If both (4.7) and (4.9) are satisfied, the 2^a users are multiplexed on the subchannel by getting $S/2^a$ slots each, otherwise only the initial 2^{a-1} users are given the $S/2^{a-1}$ slots.

The number of sharing users is doubled at each step until one of the conditions are invalidated or when the maximum number of users S is reached.

4.4.2 Rate–User Multiplexing Algorithm (Rate–Mux)

This algorithm is similar to the previous one, except that first, users are ordered according to their rate metrics. Compared to *PF–Mux*, we aim here at further increasing the cell throughput while improving PF compared to the reference *Max CSI* algorithm where each subchannel is allocated to the user with the best SNR level. In this case, condition (4.7) will hold true if the 2^a users have the same rate. This can often arise, since the discrete AMC model maps the user SNRs into 5 levels. These users are multiplexed as long as condition (4.9) holds.

4.4.3 Rule for Definite Subchannel Ordering

As it has been explained in section 3.4, the drawback of the algorithms in previous works as [30] is that subchannels are allocated in an arbitrary order. For algorithms where R_k is updated after every subchannel allocation, two different subchannel orderings can result in different user mappings. To avoid this problem here, a prioritization metric for definite subchannel selection is introduced. The first subchannel allocated for *PF-Mux* corresponds to the best PFS user among all the best PFS users in each subchannel, and so on. Similarly, the first subchannel to be chosen is the one for the best rate user among all the best rate users for *Rate-Mux*. If there are several subchannels with the same best rate, the first one is randomly picked.

4.4.4 Specification of a Reference Algorithm

Since there is no user multiplexing in *Conventional MC-PFS*, we designed another reference algorithm to ensure a fair comparison, *Conventional MC-PFS with User Multiplexing* (Mux MC-PFS), which applies *Conventional MC-PFS* slot-by-slot. Since the channel is considered constant over a frame, the instantaneous user rates are constant for the whole frame. Thus, average user rates and user PFS metrics can be determined for every slot of the frame, e. g., user allocation is calculated *in advance* for every slot of that frame. Here a different user can be allocated each slot. However, the GPF is updated on a per-frame basis, as each user perceives the PF of the allocated resources per frame.

4.5 Allocation overhead

To ensure a fair comparison, we investigate the amount of signalling needed for the DL user mapping information. We define the *net cell throughput* $\tilde{\tau}$ as the cell throughput multiplied by the ratio between the data frame duration and the frame duration plus overhead.

- **Overhead for *Conventional MC-PFS*:** the number of bits required for overhead is equal to the product between the number of bits for expressing user ID n_{ID} and the number of subchannels, M :

$$B_{CV} = M \times n_{ID}. \quad (4.10)$$

- **Overhead for *Mux MC-PFS*:** The simplest scheme for signalling the user mapping is to specify the IDs of the scheduled users for each subchannel and for each slot. The amount of bits for signalling is in this case the product

CHAPTER 4. PROPORTIONAL FAIRNESS IN A MULTI-CARRIER SYSTEM WITH MULTI-SLOT FRAMES

of the number of bits for user ID n_{ID} , the number of slots S and the number of subchannels M . But this overhead can be reduced using the fact that *the channel quality is assumed constant over a frame*. Thus, even if users are determined per slot, in a subchannel, *the slots allocated to different users can be put in any order*. Let us assume that the output of the scheduler is as depicted in Fig. 4.3(a), for a system with 4 slots, 3 subchannels and 3 users. Since the channel quality is assumed constant over the frame, *is does not matter which slots in a particular subchannel are allocated to a user, only the number of allocated slots is important*. Thus, the allocation in 4.3(a) is equivalent to the one in 4.3(b), where the user allocation is rearranged but each user is allocated the same number of slots as in 4.3(a), in each subchannel. This rearrangement enables a considerable compression of the signalling information. That is, instead of specifying a user per slot, it is enough to specify the *number of slots* for each user in a subchannel. This enables the following compression. In each subchannel, one can choose S users among K_{max} users, and any user can be selected for any slot. Thus, the number of possible allocations is equal to the number of combinations of S users among K_{max} users with repetition, which is equal to the combination of S users among $K_{max} + S - 1$, e. g., $\overline{C_{K_{max}}^S} = C_{K_{max}+S-1}^S$. With this encoding method denoted *Mux MC-PFS E1*, the number of bits needed for signalling the whole frame is

$$B_{Mux,E1} = M \times \lceil \log_2 \overline{C_{K_{max}}^S} \rceil. \quad (4.11)$$

Note that this is an information-theoretic minimum, e. g., it provides the lower bound on the amount of overhead because the exact number of bits required to represent each possible state is used. However, this bound may not be achieved by practical encoding methods, but the specific coding to approach this minimum is out of the scope of this thesis. A more common encoding scheme is to specify the allocated users and number of slots by specifying the pair $\{User\ ID, \#slots\}$ for each subchannel. By denoting $U(m)$ the number of sharing users in subchannel m , the number of bits used with this second method *Mux MC-PFS E2* becomes

$$B_{Mux,E2} = \sum_{m=1}^M U(m) \times (n_{ID} + \lceil \log_2(S) \rceil), \quad (4.12)$$

since $\lceil \log_2(S) \rceil$ is the number of bits needed for representing all the possible numbers of allocated slots.

For the signalling, we assume that QPSK modulation is used. Thus, the number of OFDM symbols needed to send this information is

$$S_{Mux,Ei} = \frac{B_{Mux,Ei}}{2 \times N}, \quad (4.13)$$

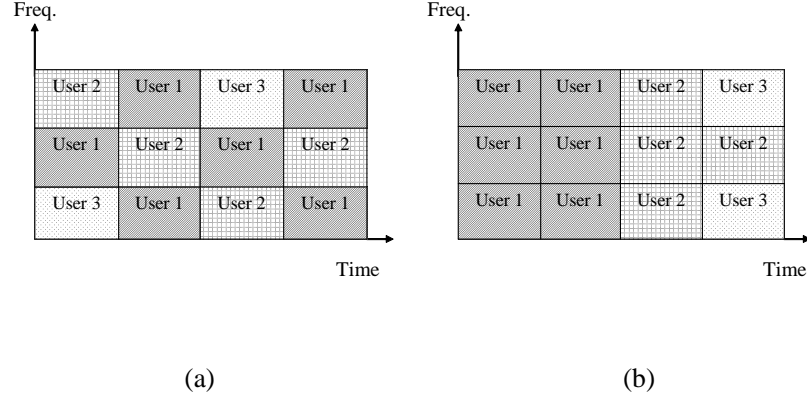


Figure 4.3: User Allocation for the *Mux MC-PFS* algorithm, with $S = 4$, $M = 3$ and $K = 3$: (a) Allocation determined by the scheduler (b) Equivalent allocation, under the assumption of a constant channel quality in a frame

and the net throughput for *Mux MC-PFS* is obtained by

$$\tilde{\tau} = \left(\frac{S}{S + S_{Mux}} \right) \times \tau, \quad (4.14)$$

where S and S_{Mux} are expressed in terms of number of OFDM symbols.

- Overhead for proposed algorithms:** In the proposed algorithms, the possible number of multiplexing users is limited since this number is doubled at each step, as explained in section 4.4.1. For example, for $S = 4$, the possible number of multiplexed users can only be 1, 2 or 4. It is further assumed that the user allocation undergoes a predetermined multiplexing pattern, as shown in Fig. 4.4, for one subchannel. By denoting U the number of multiplexed users, the pattern can be either to alternate users over slots, or to concatenate the slots for each user, as it can be seen for $U = 2$. Thus, the required information for signalling is, how many users are multiplexed, e. g., U , and which combination of U users among the total number of users, is actually allocated.

Therefore, the following scheme is used. For each subchannel, if we assume $S = 8$ here, the possible number of sharing users is 1, 2, 4 or $S = 8$. This

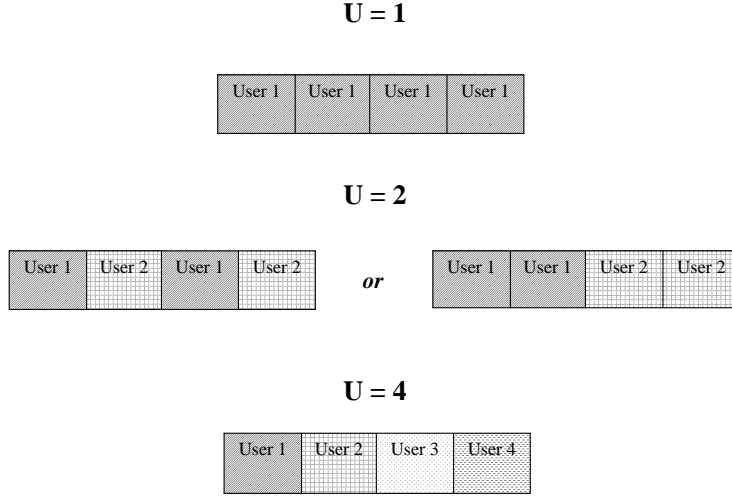


Figure 4.4: Predetermined schemes for the user multiplexing in the proposed algorithms, with $K = 4$ and $S = 4$

information is coded over 2 bits (4 possible numbers). If only one user is allocated a subchannel, there are $C_{K_{max}}^1$ possible ways to choose one user over K_{max} ; if 2 users are sharing, there are $C_{K_{max}}^2$ ways, etc. Thus, by denoting $U(m)$ the number of users multiplexed on subchannel m , the number of bits required per subchannel is $2 + \lceil \log_2 C_{K_{max}}^{U(m)} \rceil$. The total number of signalling bits for our proposed algorithms is

$$B_{Proposed} = \sum_{m=1}^M (2 + \lceil \log_2 C_{K_{max}}^{U(m)} \rceil). \quad (4.15)$$

4.6 Numerical Results

To evaluate the performance of the proposed algorithms, simulations are made over 20000 sets of channel realizations, where user locations are kept constant for a fixed number of channel realizations, then regenerated. Users are uniformly distributed in a single cell of 1000m radius. Path loss, shadowing and multipath fading channel models described in section 3.5 are used to generate each user channel. There are 64 subcarriers where BS power is equally distributed and 8 subchannels.

For the net throughput, the information overhead described in section 4.5 is considered. The number of bits for user ID is fixed to 6, allowing $2^6 = 64$ users in total. The proposed algorithms are compared with *Conventional MC-PFS* with only one user per subchannel, and *Mux MC-PFS* described in section 4.4.4 where different users can be multiplexed on each subchannel. Also, *Max CSI* provides the throughput upper bound.

As discussed in section 3.5, minimizing the window length T while achieving a high throughput translates into reducing the *maximum starvation time*, e. g., the maximum amount of time for which an individual user can be starved, i.e., not receive service. We will show that our algorithms achieve this goal, by making simulations for $T = 1680$ slots, equivalent to 1.68 seconds (with a frame duration $T_f = 8$ ms) and for a much smaller length of $T = 40$ slots. Our measure for starvation is the actual service time of a packet composed of 1000 bits, which is defined as the number of frames a user has to wait before receiving an entire packet. For each algorithm, we evaluated the maximum service time by simulations.

For $T = 40$ slots, Fig. 4.5 shows the GPF performance for the PFS algorithms, and Fig. 4.6 for *Max CSI* and *Rate-Mux*. In Fig. 4.5, the upper and lower bounds are very tight, which gives excellent approximation of the solution for optimal MC-PFS, situated between the upper bound and *Multi-Slot Rounding*. Moreover, *PF-Mux* has a near-optimal PF as it closely follows the optimal GPF. The GPF for *PF-Mux* is much higher than for *Conventional MC-PFS*, while improving the net throughput even with the overhead introduced by user multiplexing (Fig. 4.7). This proves that multiplexing can largely improve the PF performance. Moreover, *Rate-Mux* achieves a near-optimal net throughput, while achieving a much higher GPF than *Max CSI* (Fig. 4.6). Compared to the PFS algorithms, *Max CSI* and *Rate-Mux* achieve a very low GPF, which was expected as only the best SNR/rate user(s) are scheduled. The improvement in net throughput of the proposed algorithms compared to *Conventional MC-PFS* is due to the rate condition (4.7). The results show that the gain in throughput by multiplexing users having the best rates (or the best rates *among* the best PFS users for *PF-Mux*) is higher than the loss due to the allocation overhead. Moreover, while *Mux MC-PFS* achieves a near-optimal GPF, its net throughput is 20% lower than *PF-Mux* with the first encoding method *E1*, and even up to 45% lower with the second method *E2*, which is the practical case. This is due to the large amount of overhead as the user changes in each slot, which proves that simply repeating single-slot algorithms such as *Conventional MC-PFS* or the ones in [51] proposed in Chapter 3 in every slot leads to a downfall of the net throughput, which highlights the necessity of new types of algorithms such as the proposed ones, where user multiplexing is controlled.

Next, for $T = 1680$ slots, Figs. 4.8 and 4.9 depict the GPF and net throughput, respectively. We can observe that the GPF at $T = 1680$ slots increases for all algorithms, as well as the net throughput for *Conventional MC-PFS*. In this case,

CHAPTER 4. PROPORTIONAL FAIRNESS IN A MULTI-CARRIER SYSTEM WITH MULTI-SLOT FRAMES

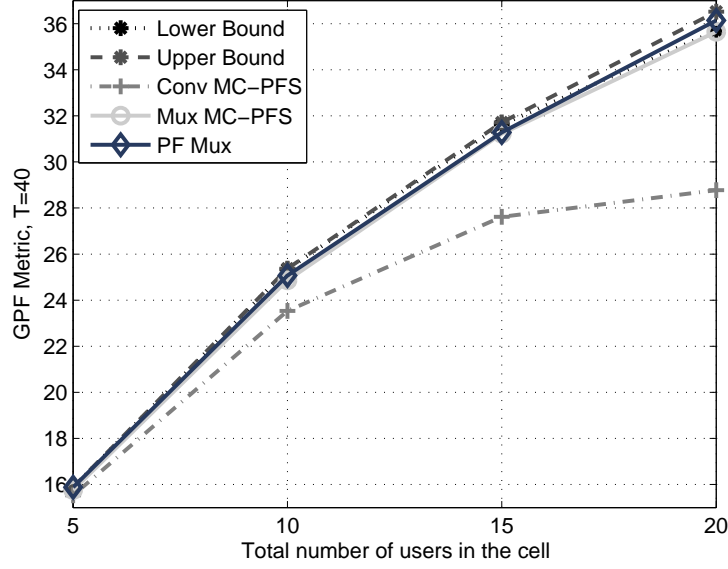


Figure 4.5: GPF metric for user multiplexing algorithms and reference algorithms, $T = 40$ slots

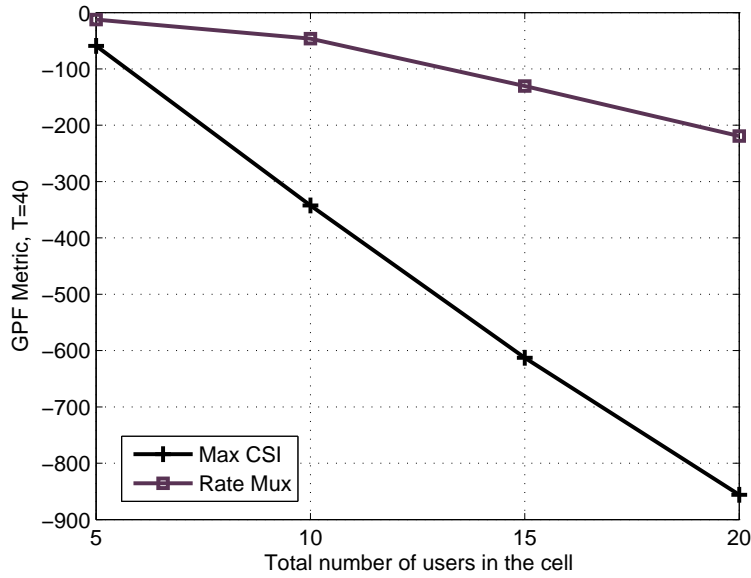


Figure 4.6: GPF metric for Rate-Mux and Max CSI, $T = 40$ slots

Conventional MC-PFS and *PF-Mux* have a similar performance. However, the maximum starvation time performance is quite different as can be seen in Table 4.1. With *Conventional MC-PFS*, the service time is 38 frames while it is only 26 frames for *PF-Mux*, which is the best performance for $T = 1680$ slots. *Mux MC-PFS* has a higher starvation time and a lower net throughput. We can think that the lower starvation time of *Mux MC-PFS* and *PF-Mux* compared to *Conventional MC-PFS* is due to user multiplexing, since several users share a frame, so that each user receives data within reduced time. Even if *Mux MC-PFS* has a very low net throughput, in average, all the users will receive one packet in a short period of time, while with *Conventional MC-PFS*, some users have to wait a longer period since only one user per subchannel is scheduled. Moreover, *Rate-Mux* achieves a lower starvation time than *Max CSI* and improves also the GPF metric. For *Max CSI*, the starvation time parameters are independent of T as the algorithm is. Compared to the PFS algorithms, the maximum starvation times of *Max CSI* and *Rate-Mux* are extremely high. This is explained by the fact that only the users with the best rate are allocated, so users with a continuously low SNR level are never served. For $T = 40$ slots, *PF-Mux* also achieves a much lower starvation time than *Conventional MC-PFS*, and *Rate-Mux* than *Max CSI*. And, as mentioned in [48], we can observe from Table 4.1 that the maximum starvation times for all algorithms increase as T increases (except *Max CSI*, since it is independent of T).

Table 4.1: Maximum starvation times for $T = 40$ and $T = 1680$, in number of frames

	Max CSI	Conv MC-PFS	Mux MC-PFS	PF-Mux	Rate-Mux
$T = 40$ slots	13500	27	7	7	1800
$T = 1680$ slots	13500	38	31	26	4800

To summarize, *PF-Mux* is able to achieve with $T = 40$ slots, the same GPF and net throughput performance as *Conventional MC-PFS* with $T = 1680$ slots, while reducing the maximum starvation time by more than 80%. Compared to *Mux MC-PFS*, *PF-Mux* tremendously improves the net throughput. Moreover, *Rate-Mux* achieves a much higher GPF performance than *Max CSI* with a lower maximum starvation time, for both values of T , while achieving a near-optimal throughput.

Finally, we have evaluated each algorithms' complexity orders, which can be found in Table 4.2. For computing the upper bound (4.4) and the Rounding algorithm, we implemented a primal-dual interior-point algorithm [46], that solves the

CHAPTER 4. PROPORTIONAL FAIRNESS IN A MULTI-CARRIER
SYSTEM WITH MULTI-SLOT FRAMES

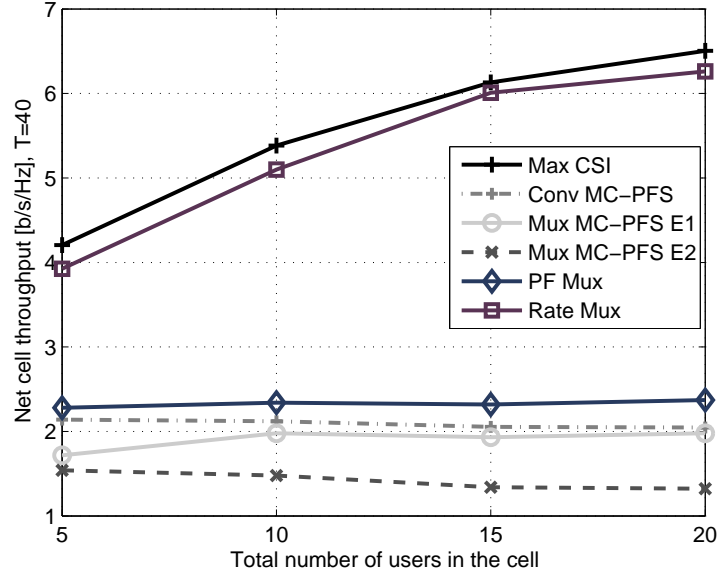


Figure 4.7: Net cell throughput for user multiplexing algorithms and reference algorithms, $T = 40$ slots

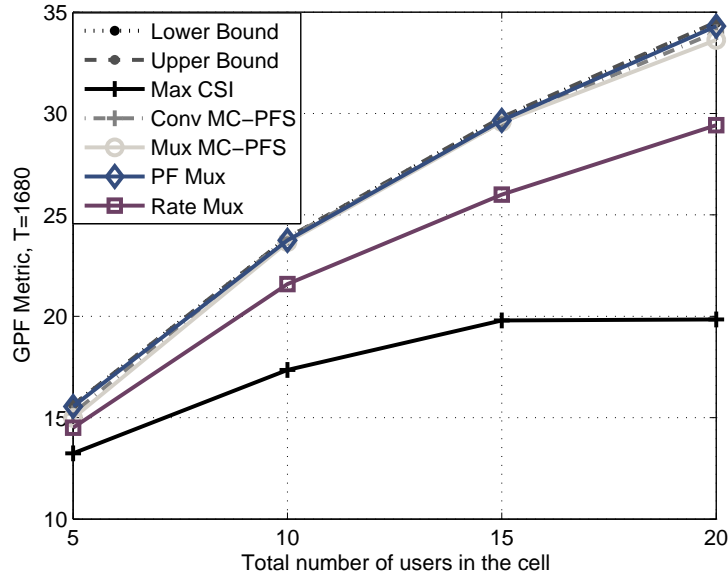


Figure 4.8: GPF metric for user multiplexing algorithms and reference algorithms, $T = 1680$ slots

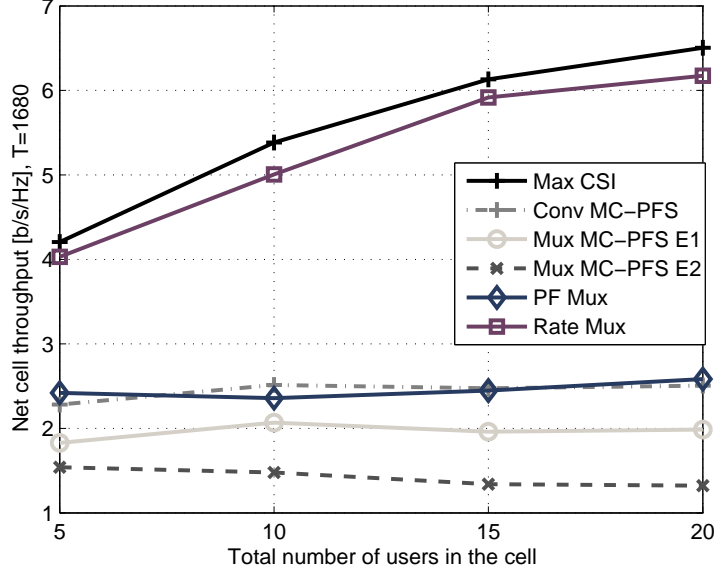


Figure 4.9: Net cell throughput for user multiplexing algorithms and reference algorithms, $T = 1680$ slots

optimality or KKT conditions using Newton's method. The most expensive step of the algorithm is solving a linear set of equations, which results in a complexity order of $O(((K+1)N)^3)$. In our proposed algorithms, the complexity is dominated by the rate and fairness conditions, (4.8) and (4.9), and the subchannel selection, which includes the user ordering by PFS or rate metric. In the worst case, the allocation is made per subcarrier, so the complexity of (4.8) is $O(N)$. In (4.9), there are $O(2(n+K))$ operations where n is the current subcarrier. Since (4.9) is repeated for each subcarrier, we have in total $O(2\sum_{n=1}^N(n+K)) = O(2NK + N^2 + N)$ operations. For the subcarrier selection, we have in the first step KN operations to determine the user PFS metrics, then $K(N-1)$, etc...until K in the last subcarrier. In total, this amounts to $O(\sum_{n=1}^N K(N-n+1)) = O(KN^2/2 + NK/2)$ operations. Since the number of steps is a constant smaller than $\log_2(S) = 3$, we obtain the complexity shown in Table 4.2. For *Conventional MC-PFS*, there are $O(KN)$ operations to determine each user PFS metric on each subcarrier. *Max CSI* has the same complexity. For *Mux MC-PFS*, this operation is repeated S times, for each slot. Compared to *Multi-Slot Rounding* which grows as $O(((K+1)N)^3)$, our proposed algorithms have 3 orders of magnitude lower complexity, while achieving a near-optimal proportional fairness for *PF-Mux*. Compared to the reference algorithms, the proposed algorithms have only around N times higher complexity while jointly improving proportional fairness, throughput and maximum starva-

tion time.

Table 4.2: Algorithms Complexity

Multi-Slot Rounding	$O(((K+1)N)^3)$
Proposed	$O((K/2+1)N^2 + (3/2K+2)N)$
Conv MC-PFS	$O(KN)$
Mux MC-PFS	$O(SKN)$
Max CSI	$O(KN)$

4.7 Summary

We have investigated the conditions for achieving PF for a wireless MC system, where a time frame is composed of multi-slots. The algorithm for obtaining the optimal solution is a prohibitively complex combinatorial problem. Instead, we have derived its upper and lower bounds by means of convex relaxation and rounding, respectively. As the rounding method is still highly complex, we have designed two heuristic algorithms where several users can be multiplexed on a same subcarrier in a time-division manner. Simulation results have validated the bounding approach which produced very tight bounds, thereby providing an accurate solution for optimal MC-PFS. Moreover, the proposed algorithms largely improve the throughput, PF and starvation time compared to reference algorithms, even with the increased signalling overhead due to user multiplexing.

*CHAPTER 4. PROPORTIONAL FAIRNESS IN A MULTI-CARRIER
SYSTEM WITH MULTI-SLOT FRAMES*

Chapter 5

Adaptive Provision of CSI Feedback in OFDMA Systems

5.1 Introduction

As already mentioned in section 2.6, with the knowledge of the user CSI at the Base Station (BS), OFDMA offers the possibility to exploit the multi-user diversity gain with a fine per-subcarrier granularity, which can maximize cell throughput. However, the amount of CSI grows linearly with the number of users and subcarriers, which can result in a prohibitively high amount of feedback in the UL channel. Usually, the following assumptions are made to limit the amount of feedback: feedback only the corresponding AMC levels instead of the SINR values, or feedback on a per-subchannel basis rather than per-subcarrier basis, where a subchannel is defined as a group of contiguous subcarriers. [34] and [35] showed that not all the feedback is required to keep the multi-user diversity gain in scheduling. Thus, the AMC level of a subchannel should be reported only if it is higher than a predefined threshold, since this subchannel has a higher probability to be scheduled. In [34] and [35], it was shown that this threshold is dependent on the number of users in the system. In this chapter, we further extend those ideas by adapting the amount of feedback depending on the scheduler considered. An optimal feedback should therefore adapt to the current channel conditions, number of users and the type of scheduler. This variable amount of feedback should be determined by the BS and requested to the users depending on the time-varying channel conditions. The key question here is, how the BS should choose the AMC levels to be reported in function of the aforementioned parameters?

In this chapter, we first show that the probability of scheduling a subchannel depends on the number of users and the scheduler. Then, we propose an adaptive feedback encoding method which can optimize the amount of feedback according

to the variable amount of CSI requested by the BS. We show that this scheme can achieve a significant feedback reduction while keeping a good scheduling performance by means of computer simulation.

5.2 System Model

The CSI that each user feeds back to the BS consists of the achievable discrete AMC level for each subchannel and each time frame. We can reasonably think that subchannels with higher AMC levels have a higher probability to be scheduled than those with low AMC levels. We consider the discrete AMC model described in section 2.8, where the SNR levels of each user in each subchannel are quantized by the AMC thresholds in Table 2.1, determined for uncoded M-QAM symbols. The corresponding AMC levels are numbered as shown in Table 5.1. Therefore, we consider that the BS formulates its requests to each user by sending messages of the following type: "Report the AMC level and subchannels with the α -best AMC levels", where α depends on the number of users and the type of scheduler. For example, if $\alpha = 2$, all the users report their subchannels supporting AMC levels 5 and 4 defined in Table 5.1.

In this chapter, we define the AMC level or mode of a subchannel as the level l corresponding to the SNR threshold immediately below the subchannel SNR. The CSI of a user is the group of the L AMC levels for the M subchannels, for one scheduling time frame.

Furthermore, we consider that the average user SNR over the whole bandwidth, or the corresponding average AMC level is known at the BS, even for users who don't have any subchannels to report. This is a reasonable assumption since the average AMC level needs to be updated only few times as it is slowly varying, which requires a very small number of feedback bits. Thus, in our scheme, when there is no CSI available for a subchannel, the user with the highest average AMC level is scheduled. By doing this, we always avoid the scheduling outage situation of [34] in the single carrier case, where the channel remains empty if nobody reports.

Table 5.1: Discrete AMC Model

Modulation	BPSK	4-QAM	16-QAM	64-QAM	256-QAM
Rate r [b/symbol]	1	2	4	6	8
AMC Level l	1	2	3	4	5
SNR Threshold [dB]	-5	13.6	20.6	26.8	32.9

5.3 Amount of Useful Feedback

5.3.1 Dependency on the Number of Users

We consider that a subchannel is reported only if it can support an AMC level $\geq \tilde{l}$. We denote $\gamma_{k,m}$ the SNR of user k on subchannel m and $\gamma_{\tilde{l}}$ the SNR threshold for AMC level \tilde{l} . We define $P_{\gamma_{k,m}}(\gamma_{\tilde{l}})$ the probability that $\gamma_{k,m} \geq \gamma_{\tilde{l}}$ which is equivalent to the probability of reporting subchannel m . Assuming that all the subchannels are fading independently, the probability $P_{m,\tilde{l}}$ that at least one user will report subchannel m can be expressed as

$$P_{m,\tilde{l}} = 1 - \prod_{k=1}^K (1 - P_{\gamma_{k,m}}(\gamma_{\tilde{l}})), \quad (5.1)$$

where K is the total number of users in the cell. In this function, we can see that $P_{m,\tilde{l}}$ increases with K . This is due to the multi-user diversity effect, where the more users there are, the more channel peaks the system experiences.

5.3.2 Dependency on the Scheduler

The amount of useful CSI depends on K , but it also depends on the nature of the scheduler as the scheduled subchannels are different depending on the scheduling. This is illustrated by simulating two types of schedulers: the *Maximum CSI* (Max CSI) algorithm which allocates in each subchannel the user with the highest SNR $\gamma_{k,m}$, and the *Conventional MC-PFS* algorithm presented in section 2.5.

Figs. 5.1 and 5.2 represent the occurrence of a scheduled subchannel having an AMC level l , normalized over the total number of channel realizations, where l is defined in Table 5.1, for $K = 5$ and $K = 50$ respectively. The scheduling is made here with full CSI, e. g., when the BS has the knowledge of the CSI of all the subchannels for all users. It can be seen that the pattern of scheduled modes is very different between the algorithms, and also between the number of users. The probability of scheduling subchannels with higher modes increases when $K = 50$. For example, only the highest mode $l = 5$ is scheduled for Max CSI. On the other hand, the PFS algorithm uses all the modes, with the lowest AMC level having the lowest occurrence. This is because subchannels with lower modes also need to be scheduled in order to maintain the proportional fairness. These results show that the reporting scheme should be different for different algorithms, since the probability of scheduling a particular subchannel with a certain mode is different.

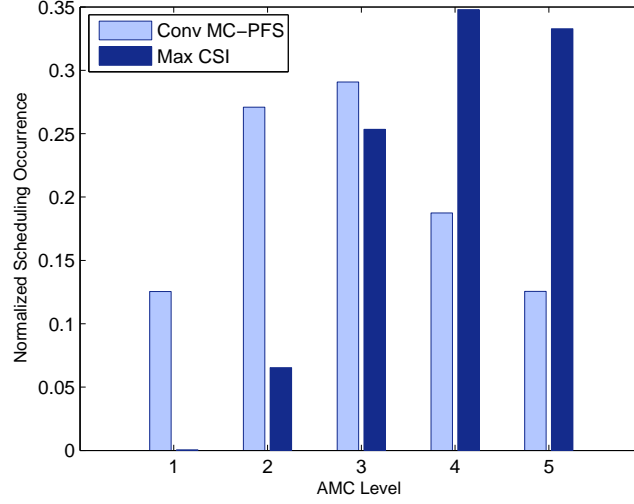


Figure 5.1: Normalized scheduling occurrence of a subchannel with mode l for $K = 5$ users, sorted from highest to lowest AMC level

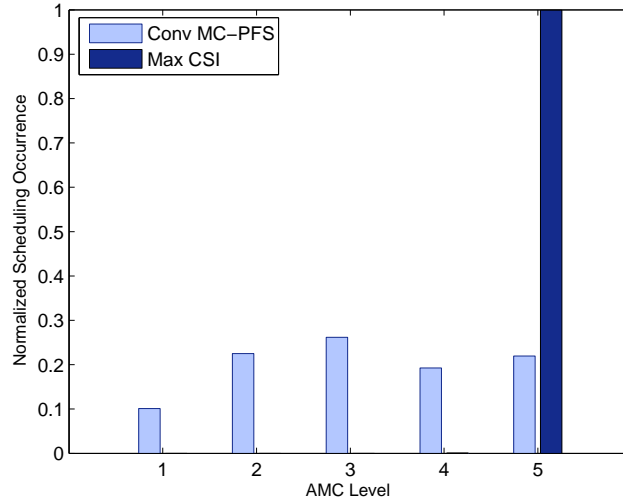


Figure 5.2: Normalized scheduling occurrence of a subchannel with mode l for $K = 50$ users, sorted from highest to lowest AMC level

5.4 CSI Feedback Encoding Schemes

5.4.1 Reference Scheme

In the conventional way, the achievable AMC level is specified for each subchannel, using $\lceil \log_2(L) \rceil$ bits per subchannel where L is the total number of AMC levels. Thus the number of feedback bits is equal to $M \times \lceil \log_2(L) \rceil$ bits.

5.4.2 Basic Schemes for Encoding one AMC Level

Here we define two basic schemes which will be used later in our overall feedback mechanism. One way of encoding a specific AMC level is to employ the *Indicator Coding* scheme, where each subchannel supporting the requested AMC level is indicated by bit 1 and the others by bit 0. A field specifying the coded AMC level is needed before encoding the subchannels. Alternatively, the *Position Coding* can be used, where the position of each subchannel having the requested level is coded. For encoding a subchannel position, $\log_2(M)$ bits are used. For example, let us assume the following OFDM channel realization with the highest available AMC levels for a user, as shown in Table 5.2.

Table 5.2: Example of achievable AMC levels

Subchannel	1	2	3	4	5	6	7	8
AMC Level	4	5	2	3	1	5	4	5

Table 5.3: Example of Feedback Message with Indicator Coding

AMC Level	Encoding
101	01000101

Table 5.4: Example of Feedback Message with Position Coding

AMC Level	Encoding
101	010110111

To report level 5 with Indicator Coding, the message becomes as shown in Table 5.3.

With Position Coding, we encode the subchannel positions corresponding to AMC level 5, namely subchannels 2, 6 and 8, so that the message becomes as in Table 5.4.

These schemes can be used when only one AMC level needs to be reported. In the following, we present encoding schemes which can be used to report several AMC levels at a time.

5.4.3 Proposed Grouped Encoding Scheme

Table 5.5: Example of Feedback Message with Grouped Coding

AMC Level	Encoding
11110	110111100101011110111

In the *Grouped Encoding* method, a predefined code is attributed to each AMC level requested. Each code should contain the smallest number of bits as possible, by using compressed codes such as Huffman codes. This scheme can be optimized by using the shortest Huffman codes for the most frequent AMC level, e. g., corresponding to the average SNR level. Moreover, the less requested levels there are, the smaller the number of bits to specify each level. Here, if α levels are reported, we use $\lceil \log_2(\alpha) \rceil$ bits to encode each of them, preceded by bit 1 to indicate that there is a report. Subchannels without report are indicated by bit 0. For example, if the minimum mode requested is $l = 2$, then $l = 2$ is coded by 100, $l = 3$ by 101, $l = 4$ by 110 and $l = 5$ by 111. We suppose that AMC levels 5, 4, 3 and 2 should be reported for the OFDM channel realization shown in table 5.2. In the beginning, the requested levels are coded by indicating them by 1 among the L levels. Then, with the codes defined above, the feedback message obtained with Grouped Encoding is shown in Table 5.5.

Since the codes are predefined depending on the number of AMC levels to be reported, this message can be uniquely decoded.

5.4.4 Proposed Differential Encoding Scheme

We propose our Differential Encoding method which can be used alternatively to the Grouped Encoding. In this scheme, we use the basic *Indicator Coding* scheme. We consider $L = 5$ AMC levels and $M = 8$ subchannels. We suppose that AMC levels 5, 4 and 3 should be reported for the same OFDM channel realization of Table 5.2.

The requested AMC levels (here 5, 4 and 3) are coded in the first field by indicating by 1 the reported AMC levels among the L levels. Then, all the subchannels with one of the required AMC levels (5, 4 or 3) are coded 1. This gives the following coding

$$11100 \quad 11010111 \quad (5.2)$$

CHAPTER 5. ADAPTIVE PROVISION OF CSI FEEDBACK IN OFDMA SYSTEMS

Next, the best level (5) is reported by indicating the corresponding subchannels *among the subchannels already marked by 1*. Thus, the subchannels marked 0, namely subchannels 3 and 5 are ignored and the subchannels with level 5 are reported only among the remaining subchannels. This gives the following coding

$$010101 \quad (5.3)$$

We repeat the same mechanism for reporting AMC level 4, this time by ignoring the bits marked 1 in the previous field of (5.3) as they contain AMC level 5. We obtain the coding below

$$101 \quad (5.4)$$

Finally, subchannels with AMC level 3 don't need to be coded as shown in the reconstruction process, described below. The total feedback message with Differential Coding is given in Table 5.6.

Table 5.6: Example of Feedback Message with Differential Encoding

AMC Level	Field 1	Field 2	Field 3
11100	11010111	010101	101

After the BS receives this message, the recovery of the information is made as follows.

- First, the subchannels bearing an information are recovered from Field 1 and marked (*), and the ones with the best AMC level 5 from Field 2. This gives the message of Table 5.7.

Table 5.7: Reconstruction 1

Subchannel	1	2	3	4	5	6	7	8
AMC Level	*	5	0	*	0	5	*	5

- To reconstruct AMC level 4, we just skip the subchannels marked 0 and the ones with AMC level 5. We reconstruct by filling the places with "*" with the indicators of Field 3, which gives the message of Table 5.8.

Table 5.8: Reconstruction 2

Subchannel	1	2	3	4	5	6	7	8
AMC Level	4	5	0	*	0	5	4	5

Table 5.9: Reconstruction 3

Subchannel	1	2	3	4	5	6	7	8
AMC Level	4	5	0	3	0	5	4	5

- Finally, the remaining subchannels marked * are for the last remaining AMC level, $l = 3$. We obtain the message of Table 5.9.

We can see that with this method, only the useful information is coded, and with a small number of bits.

5.4.5 Complete Process for Feedback Encoding

To summarize, the complete feedback message is shown in Table 5.10

- If only one AMC level is reported: use Position or Indicator Coding
- If several AMC levels are reported: use Differential or Grouped Coding. Thus, only one bit is needed to specify the encoding scheme, as the number of reported AMC levels is known from the first field.

In each case, the encoding scheme which requires the least number of bits is selected in order to minimize the amount of feedback. The structure of the message makes it easily adaptable to the varying AMC level requests and provides an optimized amount of feedback depending on the conditions.

5.5 Numerical results

We have evaluated the performance of Max CSI algorithm and Conventional MC-PFS algorithm for different number of reported AMC levels. However, the proposed methods can be potentially used with any other type of schedulers, not limited to these two. The simulation is made over 20000 sets of channel realizations,

Table 5.10: Complete Feedback Message

Field length	Content
L	String which indicates which AMC levels are reported
1	Indicates whether Position/Indicator or Grouped/Differential Coding is used
Variable	Encoding

each set of channels consisting of independent user channels. We have considered a single cell with a 1000 m radius where users are uniformly distributed. Path loss model for urban areas and log-normal shadowing with 10dB standard deviation are used. Multipath Rayleigh fading channel models with exponential power delay profile from [50] are used to generate each user channel. 32 subcarriers and 8 subchannels are considered so that 4 contiguous subcarriers are contained in each subchannel, and the number of users is varied from 5 to 30. The algorithms are run on a per subchannel basis where the CSI is the average SNR of the 4 subcarriers contained in each subchannel. In the simulations, we compare the performance of Max CSI algorithm and Conventional MC-PFS algorithm for different amount of CSI feedback with our proposed encoding scheme. The performance metric for Max CSI is the cell throughput averaged over the simulation time and for Conventional MC-PFS, the GPF parameter defined in section 2.5 which measures the level of proportional fairness, also averaged over the simulation time.

In Fig. 5.3 is represented the achieved cell throughput with Max CSI. Fig. 5.5 shows the percentage of utilization of the UL channel by the CSI feedback, assuming an UL capacity of 150 kbps. We can see that, for a low number of users, when only the three best AMC levels are reported, the achieved throughput is very close to the one achieved in the full CSI case. By reporting the three best levels, the amount of CSI feedback is considerably reduced as shown in Fig. 5.5, up to 45% in average. Moreover, as pointed out in section 5.3, the performance for the reduced feedback cases grows with the number of users, due to the multi-user diversity effect. Thus, for $K \leq 20$, the 3 best AMC levels can be reported, but this can be reduced to the 2 best levels for $20 \leq K \leq 30$, up to reporting only the best AMC level for $K \geq 30$ which leads to a feedback reduction of 90%.

In Fig. 5.4 is shown the GPF metric (see (2.8)) for the Conventional MC-PFS algorithm. We can see that the GPF rapidly degrades as the number of reported AMC levels is reduced. This is due to the fact that with a lower amount of feedback, only the good subchannels are known and the radio resource can not be distributed to users having a poor channel condition. Thus, the level of proportional fairness is not properly tuned. For users experiencing this unfairness, the average rates R_k decrease until becoming smaller than one, which gives a negative logarithm. By summing up all those negative values, a very small GPF value arises as shown in Fig. 5.4. The GPF gets smaller and smaller as the number of users K increases, since the number of negative average rates increases with K . For PFS algorithm, the four best AMC levels are needed in order to ensure a good proportional fairness, namely $l = 2$ to $l = 5$. But with our scheme, a 20% feedback reduction can still be achieved compared to full CSI feedback.

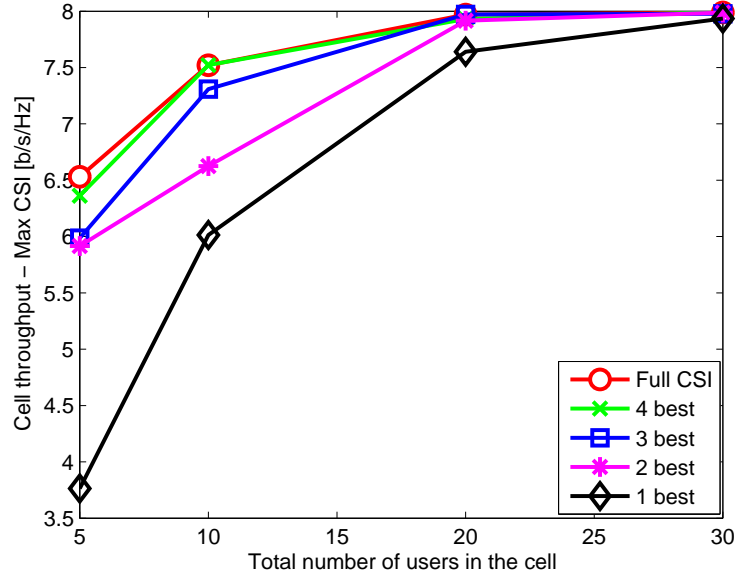


Figure 5.3: Cell throughput for Max CSI algorithm, for different levels of reported AMC levels

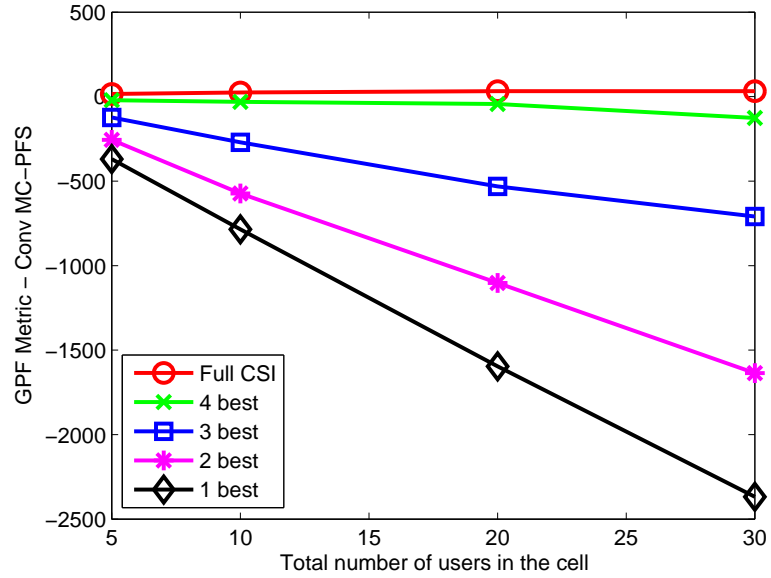


Figure 5.4: GPF metric for PFS algorithm, for different levels of reported AMC levels

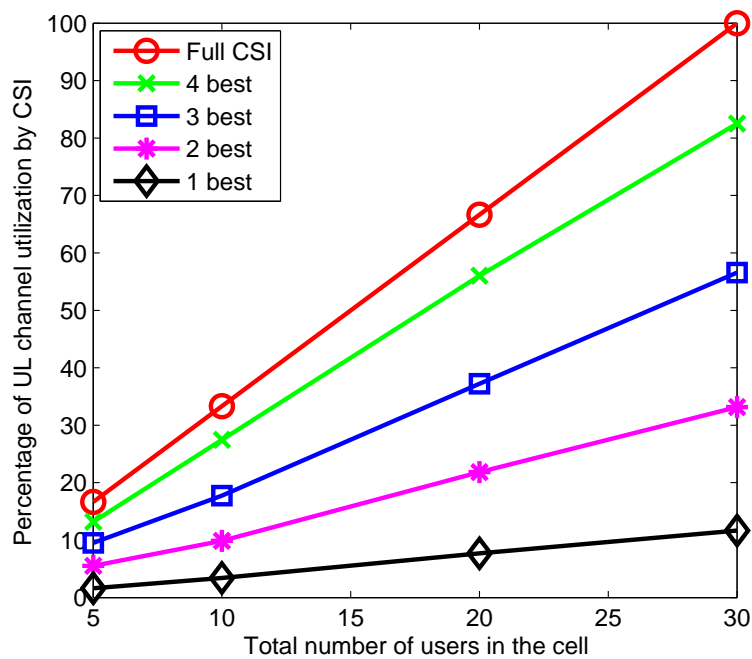


Figure 5.5: Amount of UL CSI Feedback with the proposed scheme, for different levels of reported AMC levels

5.6 Summary

In this chapter, we have studied the problem of CSI feedback in a cellular OFDMA system. Depending on the number of users in the cell and the type of scheduler being used, the useful information about the AMC levels of the subchannels varies. In this context, we have proposed a method for CSI provision which can adapt to these variable requirements and where the message length can be optimized. Simulation results have shown that a significant feedback reduction was achieved with our proposed scheme, while keeping a good scheduling performance. For Max CSI algorithm, the feedback can be reduced up to 45% for less than 20 users, 70% between 20 and 30 users and up to 90% for more users. For PFS algorithm, AMC levels from $l = 2$ need to be reported in order to keep a good proportional fairness level, but still the feedback can be reduced up to 20% with our scheme. These results have clarified the fact that the number of reported AMC levels needs to be adapted to each scheduler and to the number of users in the cell, for which our proposed feedback method provides an efficient signalling. Note that the proposed methods are generic and can be adapted to various types of algorithms.

Part II

Radio Resource Allocation and Scheduling in Relay System

Chapter 6

Throughput–Guaranteed Resource Allocation Algorithms for Relay–aided Cellular OFDMA System

6.1 Introduction and Motivation

Relays are highly considered to become an integral part of the next generation system, due to their potential for increasing the cell coverage with low costs. However, there are very few works concerning resource allocation algorithms for a relay–aided cellular system, with the assumption of the OFDMA technology. It was already seen in the previous chapters that allocation strategies are very complex in a MC system, due to the multiple degrees of freedom. The complexity of the problem intensifies with the adjunction of relays, as the number of possible links increases as well. Finding the best allocation strategy among all the possible ones, yet with low computational complexity, is a very difficult problem when so many parameters are involved.

Another point is that, the main goal of introducing relays is to provide service to the users which are not efficiently served by the BS, therefore to decrease the system outage, a viewpoint shared by the IEEE 802.16's Relay Task Group [21] (see section 2.7). Otherwise, had the goal been to maximize the system throughput, then the users with the best link conditions should be served directly by the BS, for example using the Max CSI algorithm which allocates the user with the best CSI. Thus, our main goal is to design practical allocation schemes for an OFDMA based system with relays, that focus on decreasing the outage compared to reference algorithms. In addition, the aim is to provide a good throughput per-

formance, while keeping a low complexity and minimizing the required amount of CSI.

In this chapter, after presenting the system model and path selection method, the problem of resource allocation for the single RS case is considered and algorithms for subchannel/time allocation are proposed. In the second part, the algorithms are generalized to the multiple RS case, where different numbers of relays are considered. The CSI reduction obtained with our proposed algorithms is shown. The algorithms are then evaluated and compared with reference algorithms.

6.2 System Model

As in the previous chapters, DL transmissions are considered, from a BS to MS or RS in a single cell, where users feed back to the BS their CSI on every subchannel. The relays always decode the packets and remodulate them before forwarding. The discrete AMC model of Table 2.1 is used, where the modulation is adapted for each user and each subchannel. Each RS is half-duplex, e. g., it can not transmit on one channel and receive on another at the same time. In this chapter, it is assumed that each RS is deployed to be in a Line Of Sight (LOS) condition with the BS, which ensures high channel quality on average over time and frequency. Thus, in the allocation, one crucial assumption is that all the BS-RS subchannels for one relay are allocated with the same rate, corresponding to the average SNR over the subchannels.

6.2.1 Frame structure

Case of a Single Relay

This system, depicted in Fig. 6.1, is modelled as an axis between the BS and a RS, along which users are generated. The RS is placed at a distance of $0.8 \times R_c$ from the BS, where R_c is the cell radius. The frame structure for the single RS case is shown in Fig. 6.2¹. The following assumptions are taken,

- The transmissions between BS and RS are divided in time by T_{BS} and T_{RS} respectively. The portion of the frame with BS-originated transmissions is referred to as *BS-Subframe*, and the portion with RS-originated transmissions, *RS-Subframe*.

¹In practice, the frame also contains fields for preamble, mapping messages for BS and RS, and a guard interval. The algorithms are designed without considering these fields, however, they are taken into account as overhead in the evaluations.

- Inside the BS-Subframe, BS-MS (e. g., direct) transmissions, and BS-RS (e. g., feeder) transmissions are allocated different subchannels.
- T_{BS} , T_{RS} can be adapted per frame, the basic assumption being equal time division.

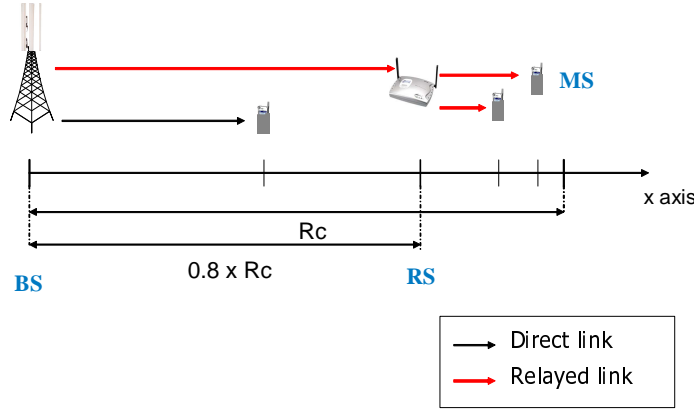


Figure 6.1: System Model for the case with Single Relay

The algorithms are optimized for the two-hop scenario, which is the most typical case and gives a good compromise between link improvement and delay [7]. With this frame structure, the packets of a relayed user queued at the BS require at least two frames to be received: in the first frame packets are sent from BS to RS; in the second frame, from RS to MS. That is, packets sent to the RS in a frame can not be immediately forwarded due to hardware limitations, since the RS needs to store and then process the newly incoming packets. Such an operation is commonly adopted in the emerging standards that use relay-based extension of the cellular systems [21] [52]. But even in the case where the RS can send the packets in the same frame, the basic idea of our algorithms can be applied straightforwardly (for example, by stacking the new packets in each subchannel if they belong to a MS already scheduled).

Case of Multiple Relays

In this second case, the BS is surrounded by I equidistant RS, as depicted in Fig. 6.3 for $I = 6$. Two types of frame division between the different relays can be considered, e. g., either in time or in frequency.

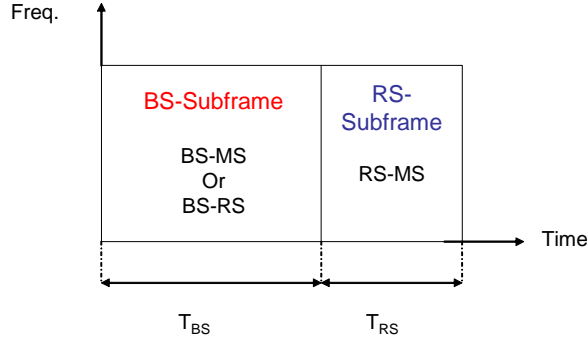


Figure 6.2: Frame structure, case of single relay

- Time division: each RS is served sequentially in time. Since the interference between the diametrically opposed relays in the cell, e.g., the pairs (RS_1, RS_4) , (RS_2, RS_5) and (RS_3, RS_6) for $I = 6$ can be assumed to be low, the structure in Fig. 6.4 will be considered, where the frequencies are reused between 2 opposite RS.
- Frequency division: each RS is assigned a frequency region as shown in Fig. 6.5, with frequency reuse. Actually, this scheme may be problematic for practical use. That is, for decoding the OFDM signal, a MS has to receive the signals at all the subcarriers at the same time. This is rendered difficult since the signals arriving at the MS are not synchronized as they come from different relays, depending on the frequency region. Moreover, there may be some frequency offset, due to the different clocks belonging to the different relays.

Due to the difficulties involved with frequency division, only time division will be considered here.

6.2.2 Path Selection

The resource allocation essentially includes two degrees of freedom: the path selection, where a user is attached either to the direct or a relayed link, second, the

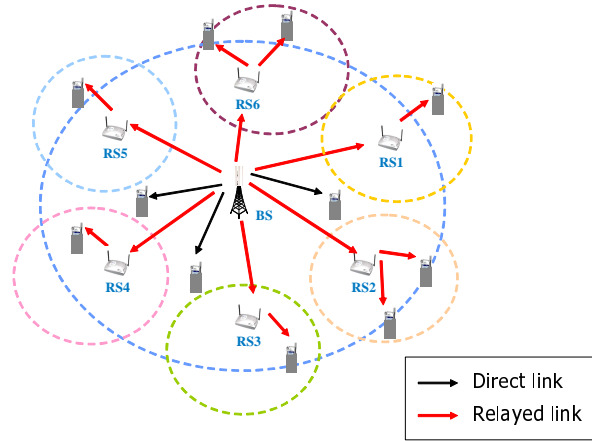


Figure 6.3: Cellular System with Multiple Relays

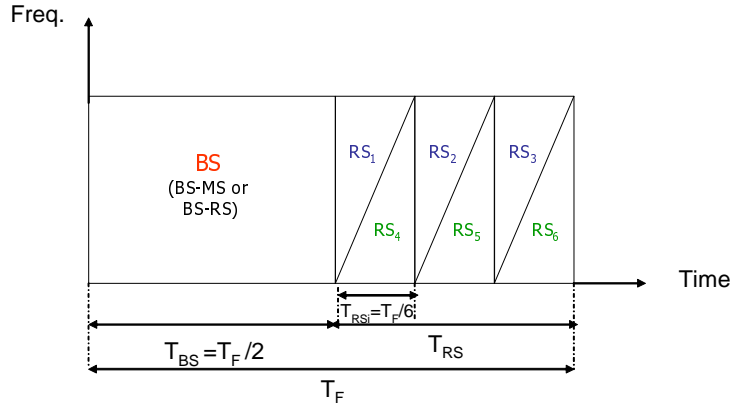


Figure 6.4: Frame Structure with Parallel RS Transmissions, time division for (MRPA) Algorithm

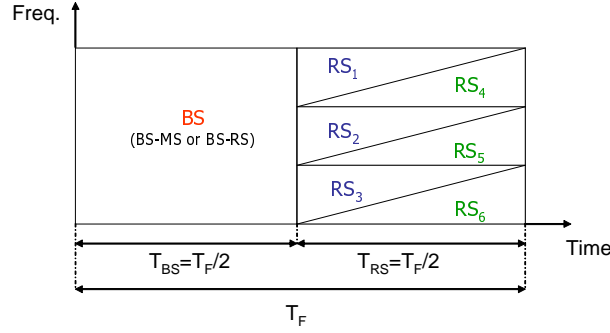


Figure 6.5: Frame Structure with Parallel RS Transmissions, frequency division

subchannel/time allocation. The optimal allocation requires the joint optimization of these two degrees of freedom by the BS. Since our goal is to provide low complexity algorithms with a reduced amount of CSI, we consider that the path selection is performed first, based on the long-term average user SNRs, followed by the resource allocation. The path selection can be done either by the BS, provided that it has the knowledge of the long-term average user SNRs over all links; or by the users themselves, who feedback the information about their chosen path to the BS, and, if a relayed path has been chosen, to that relay.

The choice between the direct or relayed paths is made as follows. Since the data for relayed users takes at least two frames to be delivered, whereas the data for direct users takes only one frame, a user is linked to the relay only if

$$\bar{r}_k^{RS-MS} \geq 2 \times \bar{r}_k^{BS-MS}, \quad (6.1)$$

and to the BS otherwise, where \bar{r}_k^l denotes the achievable discrete AMC rate averaged over time and frequency for user k on link l . In the case of multiple relays, the direct rate is compared with the rate of the best relay, for that user. The long-term channel quality for each user is periodically probed, and the path selection is renewed if there is a change. Clearly, one may argue that the average rate is not a good estimate as the user is usually allocated to a channel which is better than the average. However, a more precise estimation, such as basing the path selection on the instantaneous SNR, would involve switching the users' paths for every frame,

which is quite complex and may not be effective here, as users are considered to be static. On the other hand, the adopted estimation method offers an efficient way of making the path decision, with low complexity.

6.3 Proposed Resource Allocation Algorithms for the Case of Single Relay

The target is to design low complexity algorithms with good throughput and outage performance, while minimizing the required CSI. To achieve these goals, different approaches have been introduced. First, the allocations of RS-subframe and BS-subframe are decoupled: *the relay makes its own tentative allocation of the RS-subframe* and then, informs the BS about it. After that the BS performs the allocation of the BS-subframe and conducts optimization by considering the whole frame. This is a *RS-aided centralized* mechanism where the relay is assumed to be capable of performing its own allocation, where only basic operations are required. This decoupling between the BS and the relay enables a complexity decrease, since each subframe allocation is performed separately, but also a significant CSI reduction, as it will be shown in section 6.6. Moreover, the packets to be sent on the BS-RS link are strategically chosen: it is the relay which selects the users for which packets should be forwarded from the BS and sends this request to the BS. The selected users are the ones which have the best scheduling metric (proportional to the channel quality) but do not have any packets at the relay queue. Since low mobility users are considered, it can be reasonably assumed that a high channel quality for a certain user is likely to be kept for the next frame. Thus, with such a "just-in-time" request for user packets, it is highly probable that these packets will be delivered to destination in the following frame. The idea behind this is that, if the BS forwarded all the packets for the relayed users as they arrive in the BS queue, there would be less resource available for the direct users, as the BS-subframe is shared between the direct users and the BS-RS link. This situation becomes critical, especially in the heavily loaded case. In other words, to avoid penalizing the direct users, we optimize the number of subchannels allocated to the BS-RS link, by scheduling only the packets which will be surely delivered over the following frames. The effectiveness of this scheme will be shown later in section 6.3.3.

Details of the proposed two algorithms are presented below. In the first one, the subchannel allocation is made with the equal time division, $T_{BS} = T_{RS} = T_F/2$. The optimization of the time division is performed by an iterative approach in the second algorithm. However, the whole frame is allocated to the BS-subframe, e. g., $T_{BS} = T_F$ and $T_{RS} = 0$ in two cases: first, if all users have selected the

direct path and no one is relayed, and second, if there are no packets in the queues of the relayed users for a particular frame. In the first case, all subchannels of the frame are allocated to the direct users, but in the second one, a part of the subchannels are allocated to the BS-RS links for providing packets to the queues at the relay, following the same algorithm described below. All the proposed algorithms for relay-aided cellular system allocated only one user per subchannel, e. g., user multiplexing is not considered. However, the techniques introduced here can be combined with the ideas of user multiplexing to further boost the system performance.

6.3.1 Fixed Time Division (FTD) Algorithm

In this chapter, users are assumed to have a minimum rate requirement denoted R . Therefore, the metric used here, denoted ϕ -metric, is derived from the PFS metric but favors the users whose past average rates are furthest from the required rate R .

1. *Allocation of RS-subframe by RS:* In each subchannel n , relayed users are sorted in the order of best $\phi_{k,n}$, defined as

$$\phi_{k,n} = \frac{r_{k,n}}{\frac{R'_k(t)}{R}}, \quad (6.2)$$

where R is the minimum data rate requirement. $R'_k(t)$ is the past average rate allocated to user k at frame t over the averaging time window of T frames and is updated after every frame allocation as in (2.6).

The idea behind the ϕ -metric is to prioritize users whose achieved throughput is low compared to their rate requirement, in order to decrease the outage probability. At the same time, the users with higher instantaneous CSI are prioritized, which increases the achieved throughput. The algorithms using ϕ will be referred to as *throughput-guaranteed* algorithms, since they strive to allocate to each user a rate that best matches R . The user with highest $\phi_{k,n}$ and with packets queued at RS is allocated n .

The users having a higher $\phi_{k,n}$ than the allocated one but without packets queued at RS are represented by the set U_{Req} . For these users, the RS requests the BS to send their packets in the BS-subframe.

2. *RS sends request message to BS:* This message contains
 - the IDs of the users in U_{Req} for which packets are requested,
 - the order of these users in terms of ϕ -metric, e. g., the maximum $\phi_{k,n}$ over all n for user k in U_{Req} (needed to determine the priority of the packets sent on the BS-RS link, see Step 3)

- the value of $\bar{\phi}_{max}$, defined as the maximum value of the ϕ -metric for the relayed users in U_{Req} , averaged over the subchannels (see Step 3).
3. *Allocation of BS-subframe made by BS, based on the requests by RS:* BS allocates (tentatively) each subchannel to the best direct user. The final allocation involves the allocation of BS-RS subchannels, and is conducted as follows. If U_{Req} is non empty, the scheduler calculates the number of subchannels n_{BR} required to send all the packets queued at the BS of the users in U_{Req} . As mentioned in section 6.2, all the BS-RS subchannels are allocated with the same rate, corresponding to their average SNR level. Thus, any n_{BR} subchannels among all N subchannels can be chosen for the BS-RS transmission. To ensure a fair distribution of subchannels between the direct users and the BS-RS links, a criteria based on ϕ -metrics is introduced as follows: for each user in U_{Req} , the average ϕ -metric over the RS-MS subchannels is determined, and the maximum average ϕ -metric is denoted $\bar{\phi}_{max}$. In each subchannel, the ϕ -metric of the initially allocated direct user is compared with $\bar{\phi}_{max}$, and the subchannel is allocated to the link with the highest value. This gives y subchannels allocated to the BS-RS link. But not all y subchannels may be required, so we compare y with n_{BR} ,
- (a) If $y < n_{BR}$, the y subchannels are not enough for all the packets, e.g., some remain at the BS queue. To decide which packets to send on the BS-RS link, the RS-MS users for which packets were required are ordered by ϕ (only this order needs to be fed back to BS, not the ϕ values nor the CSI since these are not needed at the BS). Packets are allocated from the best RS-MS users, until all y subchannels are filled.
 - (b) If $y > n_{BR}$, all y subchannels are not needed for the BS-RS link since there are less queued packets. Only the n_{BR} worst subchannels for direct users are allocated to the BS-RS link, and the remaining $y - n_{BR}$ subchannels to the best direct users.

The comparison between the direct users' ϕ -metrics and $\bar{\phi}_{max}$ avoids that all the subchannels are allocated to the BS-RS link for forwarding the queued packets for the relayed users. At the same time, considering the maximum of the average ϕ -metric increases the priority of the BS-RS link, as the relayed users require twice more frames than the direct users.

For the FTD algorithm where $T_{BS} = T_{RS} = T_F/2$, the cell throughput $\tau_{T_F/2}$ is determined as a function of the allocated user rates and allocated packets. We define the channel utilization metric $u_{k,n}^l$ as

$$u_{k,n}^l = \frac{\min(r_{k,n}^l \times T_{k,n}^l, q_{k,n})}{r_{k,n}^l \times T_{k,n}^l}, \quad (6.3)$$

where l is either "BM" for direct link or "RM" for relayed link. $T_{k,n}^l$ is the number of time slots allocated to user k on subchannel n , here $T_{k,n}^l = T_F/2$ for all k , all n . $r_{k,n}^l \times T_{k,n}^l$ is the capacity of user k on subchannel n (in number of packets, with an adequate packet size), and $q_{k,n}$ is the number of queued packets for user k on subchannel n . Simply, if there are enough packets to fill the whole subchannel, then $u_{k,n} = 1$, otherwise all the queued packets are allocated and $u_{k,n} < 1$ since there are less packets than the available capacity. Note that $u_{k,n}$ is time dependent since the number of allocated packets depends on the number of allocated time slots. Thus, the throughput achieved in the BS-subframe by the direct users $k \in D$ is written as

$$\tau_{BM}(T_{BS} = T_F/2) = \sum_{k \in D} \tau_{BM}^k(T_{BS} = T_F/2) = \frac{1}{T_{BS}} \sum_{k \in D} c_{k,n}^{BM} \times u_{k,n}^{BM} \times r_{k,n}^{BM} \quad (6.4)$$

where $c_{k,n}^{BM}$ is equal to 1 if user k is allocated on subchannel n in the BS-subframe and 0 otherwise. By applying the same formula for the throughput achieved in the RS-subframe $\tau_{RM}(T_{RS} = T_F/2)$, the overall throughput is written as

$$\tau_{T_F/2} = \frac{\tau_{BM}(T_{BS} = T_F/2) \times T_F/2 + \tau_{RM}(T_{RS} = T_F/2) \times T_F/2}{T_F}. \quad (6.5)$$

The throughput achieved by the feeder link (BS-RS) is not accounted for, since the data is not delivered to the users.

6.3.2 Adaptive Time Division (ATD) Algorithm

Starting from the allocation by the FTD algorithm for $T_{BS} = T_{RS} = T_F/2$, the time division can be adapted in order to increase the overall throughput. Basically, the goal of this algorithm is to balance the allocated time between BS and RS-subframes, according to the amount of packets in the BS and RS queues. By matching the time division with the proportion of the packets in each queue, the throughput should be increased. We perform an iterative optimization where the goal is to find the best time division between the BS and RS subframes, within the fixed total frame length T_F . For this optimization, the users and the number of initially allocated packets are known for each subchannel, as the outcome of the FTD algorithm. But when the time division is changed, the number of packets allocated to each user changes: for example, if the BS-subframe is increased by one slot and the RS-subframe reduced by one, the direct users are allocated an amount of packets corresponding to the additional capacity in their allocated subchannel (if there are any queued packets), and for the relayed users, the corresponding packets are removed. This optimization problem can be formulated as follows:

$$\begin{aligned}
 & \tau_{Opt} = \max_x \tau_x \\
 \text{where } & \tau_x = \frac{\tau_{BM}(T_{BS}) \times T_{BS} + \tau_{RM}(T_{RS}) \times T_{RS}}{T_F} \\
 \text{subject to } & T_{BS} = T_F/2 + x \times T_{slot}, T_{RS} = T_F/2 - x \times T_{slot} \\
 & x \in \left[-\frac{T_F/2}{T_{slot}}, \dots, +\frac{T_F/2}{T_{slot}}\right]
 \end{aligned} \tag{6.6}$$

where the variable $x \in \mathbf{Z}$. Since this is a difficult problem due to the discrete packet updates at each time adaptation, we propose the following *Adaptive Time Division* (ATD) algorithm which performs an iterative optimization. The idea is to start from the initial condition with $T_{BS} = T_{RS} = T_F/2$ and then consider the two possible cases:

1. increase T_{BS} by one slot and decrease T_{RS} by one slot, e. g., $T_{BS} = T_F/2 + T_{slot}$ and $T_{RS} = T_F/2 - T_{slot}$. The new throughput, after the updated packet allocation, becomes:

$$\tau_a = \frac{\tau_{BM}(T_F/2 + T_{slot}) \times (T_F/2 + T_{slot}) + \tau_{RM}(T_F/2 - T_{slot}) \times (T_F/2 - T_{slot})}{T_F}. \tag{6.7}$$

2. decrease T_{BS} by one slot and increase T_{RS} by one slot, and the throughput becomes:

$$\tau_b = \frac{\tau_{BM}(T_F/2 - T_{slot}) \times (T_F/2 - T_{slot}) + \tau_{RM}(T_F/2 + T_{slot}) \times (T_F/2 + T_{slot})}{T_F}. \tag{6.8}$$

Then, we compare $\tau_{T_F/2}$, τ_a and τ_b and the maximum determines the direction for the time adaptation. There are three possible cases:

- if $\tau_{T_F/2}$ is the maximum, the algorithm stops since the adaptation in either direction gives a lower throughput, so $\tau_{Opt} = \tau_{T_F/2}$.
- if the maximum is τ_a , the time division is adapted again by increasing the BS-subframe by one slot, and reducing the RS-subframe by one slot, namely $T_{BS} = T_F/2 + 2 \times T_{slot}$ and $T_{RS} = T_F/2 - 2 \times T_{slot}$. After reallocating the packets correspondingly, adding or removing the adequate number of packets from each subframe, a new value of throughput is obtained and denoted τ_{a+1} . At iteration i , we have $T_{BS} = T_F/2 + i \times T_{slot}$ and $T_{RS} = T_F/2 - i \times T_{slot}$, with throughput τ_{a+i} . The iterative search stops when: $\tau_{a+i+1} < \tau_{a+i}$, or when there are no more slots to be removed from the RS-subframe. The maximum throughput is finally obtained by: $\tau_{Opt} = \tau_{a+i}$.
- if the maximum is τ_b , the same iterative search is applied as previously, but by reducing the BS-subframe by one slot and increasing the RS-subframe by one slot, namely $T_{BS} = T_F/2 - 2 \times T_{slot}$ and $T_{RS} = T_F/2 + 2 \times T_{slot}$.

The optimal solution requires a full search over all the possible time division. This ATD algorithm is suboptimal since it restricts the search to either one direction. However, it can still be ensured that a good division is found when the search is restrained to the direction giving an increased throughput, while decreasing the algorithm complexity. This is due to the fact that the fluctuations of the throughput in each subframe is governed by the number of packets at the queues at the BS and the relay. That is, if by adding one slot in one direction, the throughput becomes lower, it is likely that it will continue to decrease if more slots are added, as there are no more packets in the queue that can fill up the additionally allocated capacity. At the same time, the number of slots initially allocated to the other subframe are removed, which implies that the packets that were allocated initially are removed accordingly, which leads to an overall decrease of throughput. Thus, this creates a situation where one subframe is attributed too much capacity which is not utilized due to a lack of queued packets, and the other subframe has too little capacity and can not allocate a sufficient number of packets. Considering this fact, it can be thought that it is unlikely that the best allocation can be found by searching in the direction where throughput is decreased.

It can be observed that the final decision is made by the BS, so an overall optimization can be achieved up to a certain degree, even for this suboptimal algorithm. Moreover, it will be shown in section 6.6 that the amount of CSI feedback is reduced compared to the optimal BS-centralized algorithm. However, the BS needs to know the queue status at the RS to perform this adaptation, which is possible (the BS can monitor the queue status at the RS since it knows how many packets were sent to the RS and how many were allocated from the RS to the MS), but it makes the ATD algorithm less practical than the FTD algorithm. However, this increased complexity at the BS may provide some performance improvements compared to the FTD algorithm.

6.3.3 Discussion on the packet requests to the BS by the relay

In the FTD and ATD algorithms, the relay requests the BS to forward packets for some chosen relayed users, e. g., with the best ϕ -metric but without packets at the RS queue. This is a more complex way as compared to the case where the BS forwards all the packets for relayed users, e. g., in the order of arrival. However, note that in such a case, the BS is oblivious about the state of the RS-MS links. For comparison, we have designed the *All Forward* (All-Fwd) algorithm which works like the FTD algorithm, except that the RS no longer requests packets but the BS forwards the packets for randomly chosen relayed users. That is, not all the packets for relayed users are always forwarded; in the BS-subframe, only the subchannels where the maximum of the average ϕ -metric of all relayed users, $\bar{\phi}_{max}$, is higher than the best direct user's ϕ -metric, $\phi_{k,n}$, are allocated to

the BS–RS link, which determines the number of relayed users' packets that can be accommodated. The simulations will show the effectiveness of our requesting scheme compared to the *All Forward* algorithm, in terms of throughput and system outage probability (see section 6.7). This is the reason why the algorithms for multiple relays presented in the following section are also based on this requesting mechanism.

6.4 Proposed Resource Allocation Algorithms for the Case of Multiple Relays

In general, there are more than one relay in a cell. We propose algorithms for allocating resources with multiple RS. The two basic options for RS allocation presented in section 6.2 are compared: first, the time division and second, the frequency division. For both cases, the initial division is $T_F/2$ for the BS–subframe and the RS–subframe. Within the RS–subframe, there is an equal time/frequency division between the RS_i –subframes, as shown in Figs. 6.4 and 6.5.

6.4.1 Multiple–RS Parallel (MRP) Algorithm

We consider the frame structure in Fig. 6.4 for $I = 6$ relays. There are 3 groups of opposite RS, the pairs (RS_1, RS_4) , (RS_2, RS_5) and (RS_3, RS_6) , each sharing one RS_i –subframe (see Fig. 6.3).

The allocation is based on the *FTD* algorithm from section 6.3.1. The difference from the single RS case is that the allocation of the RS–subframe is performed by each RS, for each RS_i –subframe, and each RS makes its own list of users, $U_{Req}^{RS_i}$, for which packets are requested. Then, the BS considers the request lists of all relays and performs the steps in 6.3.1: in each subframe, each subchannel is allocated to the user with the best $\phi_{k,n}$, defined in (6.2). For the allocation of subchannels to the BS– RS_i links, the packets for the users in $U_{Req}^{RS_i}, \forall i$ are concatenated and allocated together. Each relay decodes each received packet, and checks the destination user IDs. Only the packets intended to their own attached users are kept. Thus, the relays only need to know which subchannels are allocated to the BS–RS links, such that the signalling remains the same as in the single RS case.

6.4.2 Multiple–RS Parallel with Activation (MRPA) Algorithm

In MRP algorithm, even if no users are attached to a certain RS, the subchannels can be used by the other RS in the same subframe. However, if two RS of a

common subframe are both empty, that resource is wasted. To circumvent this loss, the concept of *Adaptive RS-Activation* is introduced.

For $I = 6, 8, 12$, we can make $I/2$ relay pairs by regrouping the diametrically opposed relays. Each group of 2 relays transmit at the same time by sharing the same RS_i -subframe, since the interference is minimized between the opposite relays. If after the path selection, 2 opposite RS in the same group are not activated (for example, RS_1 and RS_4), their corresponding subframe is removed and the overall RS-subframe is redivided in time among the remaining groups. For $I = 6$, each subframe has an initial length of $T_F/6$ but after removal of a group, it will become $T_F/4$. However, for $I = 3$ the relays cannot be paired so frequency reuse is not considered. Instead, each of the 3 equally divided RS_i -subframe is only used by one relay. If no users are attached to a relay after the path selection, the corresponding subframe is removed and the whole RS-subframe is divided among the remaining relays.

After this relay activation step, the subchannel allocation is based on the *FTD* algorithm as for the *MRP* algorithm.

6.4.3 Multiple-RS Adaptive Activation (MRAA) Algorithm

For this algorithm, we consider an initial frame where $T_{BS} = T_F/2$ and the RS-subframe is equally divided in time between all the relays, without assuming frequency reuse, resulting in a RS_i -subframe length of $T_F/2/I$. The best adaptation of the time division between the different entities would be to perform an *ATD*-like algorithm as in the single RS case, where the time given to each subframe would be optimized. However, in the multiple RS case, this becomes a multi-variable integer optimization problem, which requires a very high complexity. Therefore, we propose the following method, where the RS activation occurs in two steps: the long-term RS activation and the per-frame RS activation.

1. First, the long-term RS activation is carried out as described previously, based on the path selected by each user. The idea is to keep only the RS_i -subframe for the relays where users have been attached. For example with $I = 6$, if nobody is attached to RS_1 and RS_5 , the remaining access points are the BS and the other four RS. The RS-subframe duration is equally divided among the four RS, resulting in the frame shown in Fig. 6.6. This is a *long-term* RS activation as it is based on the path selection which depends on the long-term average channel qualities.
2. Next, the same algorithm for subchannel allocation as in the *MRPA* algorithm is performed, based on $\phi_{k,n}$. After this initial allocation, the users and packets on each subchannel and subframe are known.

3. The throughput is then optimized by removing the worst RS_i -subframes until the best throughput is obtained. This *per-frame RS activation* phase works as follows.
 - Initially, the throughput τ_0 is calculated for all the R_0 relays and the BS. The BS-subframe has an initial length of $T_{BS}(0) = T_f/2$ and each RS_i -subframe has a length of $T_{RS_i} = T_f/2/R_0$.
 - Then, the RS are sorted in the order of decreasing throughput τ_{RS_i} achieved in each RS_i -subframe. The RS with the worst throughput is removed, and the RS_i -subframe is reallocated to the BS-subframe, so that: $R_1 = R_0 - 1$ and $T_{BS}(1) = T_{BS}(0) + T_{RS_i}$. The packet allocation and the queues are updated accordingly.
 - If the new total throughput τ_i is higher than the previous one τ_{i-1} , we continue by removing the next worst RS and redistributing the frame, otherwise the previous frame configuration is kept. This algorithm stops when $\tau_{i-1} > \tau_i$ or when only the BS subframe is left.

A concern might be that a RS is always removed if this RS always achieves a low throughput, but since all RS are statistically identical, they have an equal chance to be served. The simulation confirms that this problem does not arise.

For a high number of relays, as for $I = 12$, each RS_i -subframe length $T_f/2/I$ becomes very small. If $T_f/2/I$ becomes smaller than the minimum time allocation unit of one slot T_s , I_{max} relays are randomly chosen to be allocated in this frame, where $I_{max} = T_f/2/T_s$, and the other $I - I_{max}$ relays are discarded for this frame. In each frame, a new set of I_{max} relays are randomly chosen for allocation. After this additional step, the algorithm works as described above.

Finally, the highlights of the proposed algorithms and their design features are summarized in Table 6.1.

6.5 Design of Optimized Algorithms for Performance Assessment

To assess the performance of the proposed algorithms, optimized algorithms have been designed to provide some upper/lower bounds. So far the path selection and the subchannel/time allocation have been separated in order to reduce the algorithm complexity and to avoid the feedback of the CSI of relayed users to the BS. However, to obtain the best performance, the path selection and the resource allocation should be jointly made at the BS.

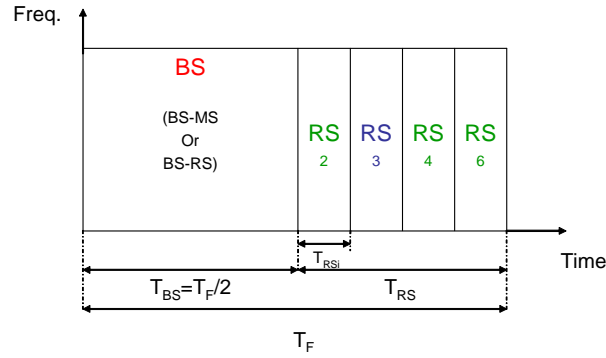


Figure 6.6: Frame Structure for the Multiple-RS Adaptive Activation (MRAA) Algorithm

Table 6.1: Summary of the Proposed Algorithms

Single Relay	FTD	Equal time division, outage based algorithm with low complexity
	ATD	Adaptive time optimization, higher throughput than FTD but higher complexity
Multiple Relays	MRP	Fixed time division with frequency reuse, outage based algorithm with low complexity
	MRPA	Fixed time division with long-term RS activation, outage based algorithm with better throughput than MRP
	MRAA	Adaptive time division with per-frame RS activation, trade-off throughput/outage

6.5.1 Upper Bound for Throughput in the case of Single Relay

The following assumptions are made to ensure an upper bound to the throughput:

- for a relayed user, all the packets coming from the BS to the RS in a frame are received by the user during the same frame,
- the time division between BS-subframe and RS-subframe is optimized *per subchannel*.

In this case, the frame becomes as in Fig. 6.7, where $T_{k,n}^{BS}$ is the time allocated for BS-RS transmission for user k on subchannel n . For direct users, $T_{k,n}^{BS} = T_F$. This algorithm gives an optimized performance since we consider unrealistic optimal assumptions: in reality, packets for relayed users need at least 2 frames to arrive to destination, and a different time division per subchannel is not feasible since a RS cannot receive and transmit at the same time on different subchannels. The real throughput upper bound would be given by taking the full buffer assumption, e. g., there are always packets to be sent for all users, and performing the Max CSI algorithm. But that results in an extremely disparate upper bound as the best user always achieves the highest rate due to the user distribution on the RS-BS axis. To obtain a tighter bound, this assumption was dropped and the following optimized scheme is adopted instead:

1. We consider K users. Set D contains the CSI of all K users for the direct link and set R contains the CSI of all K users for the relayed link. For each subchannel, we have to determine which user on which link to allocate, in order to maximize the throughput.
2. For the users in R , we determine the time division between $T_{k,n}^{RS}$ and $T_{k,n}^{BS}$. With the assumption that everything sent from BS to RS arrives at the MS during the same frame, $T_{k,n}^{RS}$ and $T_{k,n}^{BS}$ are proportional to the BS-RS and RS-MS rates on the subchannel, namely \bar{r}_{BR} and $r_{k,n}^{RM}$, and can be determined as:

$$T_{k,n}^{RS} = \frac{\bar{r}_{BR}}{\bar{r}_{BR} + r_{k,n}^{RM}} \times T_F \quad (6.9)$$

and $T_{k,n}^{BS} = T_F - T_{k,n}^{RS}$.

3. The effective capacity $\eta_{k,n}^l$ is defined as a product of the capacity (see section 6.3.1) with the channel utilization metric defined in (6.3), for each user and subchannel

$$\eta_{k,n}^l = u_{k,n}^l \times r_{k,n}^l \times T_{k,n}^l. \quad (6.10)$$

4. In each subchannel, we simultaneously order by decreasing effective capacities $\eta_{k,n}^l$ the $2K$ users from sets D and R , in order to choose the best path and the highest efficiency simultaneously. The best user, who has either link BS or RS , is allocated the subchannel. Finally, the throughput is computed.

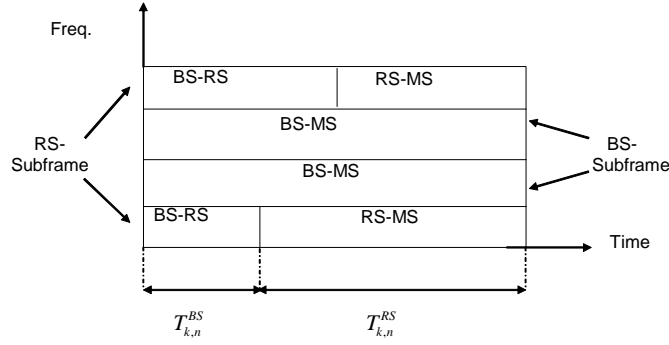


Figure 6.7: Frame structure used for the optimal algorithm

6.5.2 Upper Bound for Throughput in the case of Multiple Relays

In the case of multiple RS, compared to the previous case, the same number of users is generated over a much larger area, thereby decreasing the probability that one user will always support the highest rate 8 [b/s/Hz]. Since the throughput achieved by the Max CSI algorithm assuming full buffer is much lower than in the single RS case, this algorithm gives a suitable upper bound, referred as *RS-Max Full Buffer*. As in section 6.5.1, it is also assumed that the data sent to a RS is immediately forwarded to the user in each frame. The best user among all the direct or relayed users is scheduled in each subchannel.

6.5.3 Lower Bound for System Outage

In the case of system outage, all algorithms are evaluated with the full buffer assumption. That is, a user is considered to be in outage when he has not been

given any resource, although he has queued packets. But if he is not scheduled because he does not have any queued packets, it is not considered to be an outage event. As in section 6.5.1, we assume that the data sent to a RS is immediately forwarded to the relayed user, within each frame, and that the time is optimally divided within each subchannel between the BS–RS and RS–MS links. Since the outage (see section 6.7.2) is based on the number of users who have not received a rate higher than a reference rate R , we define our algorithm as follows: in each subchannel, all users are ordered according to their priority metric on all the access points, and the subchannel is allocated to the best user. The priority metric is ϕ defined in (6.2), but here the average user rate $R_k(t)$ is updated after every subchannel allocation, in order to obtain a finer allocation, namely

$$R_k(t)(n_c) = \frac{T-S}{T} \times R'_k(t) + \frac{1}{T} \times \sum_{n=1}^{n_c} c_{k,n} r_{k,n}(t) \quad (6.11)$$

with n_c the currently served subchannel. In the single RS case, the user on the link with the best ϕ is scheduled in each subchannel. This algorithm is denoted *Optimized Outage* algorithm.

6.6 CSI Reduction with Proposed Algorithms

A major concern for RS-aided allocation is the increased amount of CSI. An optimal algorithm performed at the BS requires:

1. the CSI for BS–MS link of all K users in the cell per subchannel per frame
2. the CSI for RS–MS link of all K users in the cell per subchannel per frame.

This results in a tremendous amount of overhead, especially as the number of RS increases. In the single RS case, for *FTD* algorithm, the required feedback is composed of:

1. the CSI of the K_D direct users at the BS, per subchannel per frame, where $K_D \leq K$
2. the CSI of the K_R relayed users at the RS, per subchannel per frame, where $K_R \leq K$
3. user IDs of users in the list U_{Req} , their order in terms of ϕ -metric, and the value of $\bar{\phi}_{max}$, sent from RS to BS.

Since $K_D + K_R = K$, the amount of information of 1) + 2) for *FTD* algorithm is equivalent to that of 1) for the optimal algorithm. It is also equal to the amount of CSI required for the Max CSI algorithm without relay. Since the number of users in U_{Req} is usually small, 3) will be reasonably small. Thus, the CSI information required by *FTD* algorithm is much lower than the one required for the optimal algorithm and slightly higher than for Max CSI. Thanks to the requests by the relay, the feedback can be minimized. For the *ATD* algorithm, in addition to the feedback of *FTD* algorithm, the CSI of the allocated relayed users is required at the BS, as it has to compute the throughput of the RS-subframe for time adaptation. But it is still much lower than the information needed for the optimal algorithm, since $K_{R,alloc} \leq K_R \leq K$ where $K_{R,alloc}$ is the number of users allocated in the RS-subframe among the relayed users.

In the multiple RS case, the information required for *MRP* algorithm is equivalent as for the *FTD* algorithm, where 2) + 3) are obtained from each RS. So the feedback reduction against the optimal algorithm is comparable to the single RS case. The feedback for *MRPA* and *MRAA* algorithms is also equivalent to the *ATD* algorithm, but from each RS. In either case, the CSI required by the proposed algorithms is much reduced compared to the optimal case, since each RS makes its own allocation and forwards only the useful information to the BS.

6.7 Numerical Results

In the simulations, a single cell with a BS and one/multiple RS is considered. Simulations are made over 100000 sets of channel realizations, where user locations are kept constant for a fixed number of channel realizations, then regenerated. Such a large number of realizations is required for the calculation of the system outage, for which a sample of the number of users in outage is taken every 100 frames (see explanations below). The cell has a 1000 m radius and the relays are placed 800 m away from the BS. The path loss model proposed in [53] is used for its simplicity and ability to adapt to the three types of links: BS-MS, BS-RS and RS-MS, where the path loss L in [dB] is given as

$$L = [51 - 8 \times \log_{10}(H_{T_x} \times H_{R_x})] \log_{10}(d) + 8.4 \times \log_{10}(H_{T_x} \times H_{R_x}) + 20 \times \log_{10}(f_c/2.2) + 14, \quad (6.12)$$

where

H_{T_x} is the antenna height at the transmitter (3.5–30 m)

H_{R_x} is the antenna height at the receiver (1.5–30 m)

d is the distance between the transmitter and receiver (200–3000 m)

f_c is the carrier frequency (2 GHz band).

The antenna heights at the BS, RS and MS were fixed to 30, 20 and 2 m, respectively, as in [54]. Log-normal shadowing with 0 dB mean and 8 dB standard deviation is assumed for the BS–MS links, and with 6 dB standard deviation for the RS–MS links. The same multipath fading channel model as described in section 3.5 is used. The BS power and RS power are fixed to 20 Watts and 5 Watts [55], respectively. Given the subframe, there is equal power distribution in each subcarrier. There are 48 subcarriers and 12 subchannels, each composed of 4 contiguous subcarriers. The frame duration is fixed to 12 ms. Packets arrive at the BS queue following a Poisson process.

6.7.1 User Generation in the cell

One of the major motivations for introducing relays is to increase the coverage of a cell, e. g., to support more users located in the outskirts of the cell. While the per-user throughput for the users located in the cell edge may increase when there are relays, it is not clear what the impact will be on the overall throughput performance, as more radio resource is used to support the additional relayed link. Therefore, it is reasonable to think that the performance of the algorithms will depend on the user distribution. If users are uniformly distributed within the cell, it may not pay off to make use of relays, as compared to the case where more users are located towards the cell edge. Thus, two cases of user distribution have been considered, the uniform distribution and the edge distribution, where users are located towards the cell edge. Depending on the single or multiple relay case, the edge distribution is modelled as follows.

1. **Case of single relay:** in this case, a simple model is used. Users are generated along the axis depicted in Fig. 6.1. If x denotes the distance from the BS, while users are generated from $x = 0$ to the cell edge $x = R_c$ in the uniform distribution, in the edge distribution, users are generated from $x = 0.4 \times R_c$ to $x = R_c$. This is not realistic since usually, there are varying densities of users depending on the location, however, it can still give an idea of the impact of the distribution on the performance metrics.
2. **Case of multiple relays:** in this case, a more realistic model derived from [56], *clustering to the edge of the cell*, is applied. The cell is divided into square bins as shown in Fig. 6.8, which are each attributed a certain probability to be selected, P_i , for $i \in [1..16]$. To generate more users towards the edge compared to the center, the bin probability increases as the distance from the

bin center to the BS increases. The bins sharing the same center to BS distance are characterized by the same bin probability. In this case, 3 regions can be defined where each region corresponds to bins of equal probability, namely:

- the central area C regrouping bins 6, 7, 10 and 11,
- the closer edge area E_1 regrouping bins 2, 3, 5, 8, 9, 12, 14 and 15,
- the further edge area E_2 regrouping bins 1, 4, 13 and 16.

Users are allowed to be generated anywhere in the whole square grouping the 16 bins, including outside the hexagonal cell. The user generation is carried out similarly as described in [56]. First, a bin is selected by sampling the Cumulative Density Function (CDF) of the bin probabilities with a uniformly distributed random number between 0 and 1. Then, a user is generated randomly within the selected bin area.

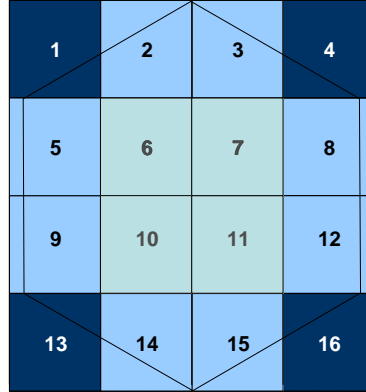


Figure 6.8: Hexagonal cell divided into bins, for generating the user distribution towards cell edge, case of multiple relays

Once users are generated either uniformly or towards the edge, each user is attached either to the BS or one of the RS, using the path selection method described in section 6.2.2. After that the time/subchannel allocation is performed. The path selection is renewed each time the average user SNR changes.

6.7.2 Performance Metrics

The system performance is characterized by means of two metrics: the goodput and the system outage. The goodput γ in [b/s/Hz], where the overhead for CSI feedback is included, is defined as

$$\gamma = \tau \times \frac{n_{data}}{n_{data} + n_{OH}}, \quad (6.13)$$

where τ is the throughput, n_{data} the number of OFDM symbols in the frame carrying data and n_{OH} the number of symbols carrying the CSI, assuming QPSK modulation.

The system outage is defined as the probability that the allocated user rates \bar{r}_k are lower than a reference rate R , where \bar{r}_k is averaged over $p = 100$ frames. The system outage probability P_{out} is expressed as

$$P_{out} = \frac{\sum_{s=1}^S K_s}{K \times S}, \quad (6.14)$$

where K_s denotes the number of users in outage for the sample s and S is the total number of samples, $K_s = \text{Card}\{k, \bar{r}_k < R\}_s$, where Card denotes the number of elements in the set. If P is the total number of frames during the simulation, $S = P/p$ since the number of users in outage is taken every p frames.

6.7.3 Single RS Case

The FTD and ATD algorithms are compared with the reference algorithms: the *Conventional MC-PFS* algorithm without relays, the *RS-Opt* algorithm which gives the throughput upper bound, and *Optimized Outage* for the system outage lower bound. The achieved goodput is evaluated for 5 to 20 users, and the number of users is fixed to 20 for the outage.

When assuming uniform user distribution, Fig. 6.9 shows that the FTD algorithm achieved a lower throughput than *Conventional MC-PFS*, while ATD algorithm has a slightly higher throughput. There is a certain gap between the goodput of ATD algorithm compared with the one of *RS-Opt* algorithm, which can be explained by the fact that *RS-Opt* algorithm only optimizes the throughput without consideration of PF. The system outage performance of the proposed algorithms is tremendously lower compared to *Conventional MC-PFS* and closely follows the outage lower bound given by *Optimized Outage*, except for $R = 200$ [kbps] as shown in Fig. 6.10. Also, the outage probability of the *RS-Opt* algorithm is unacceptably high.

However, the curve tendencies change when users are distributed towards the cell edge. Fig. 6.11 shows that both FTD and ATD algorithms drastically improve

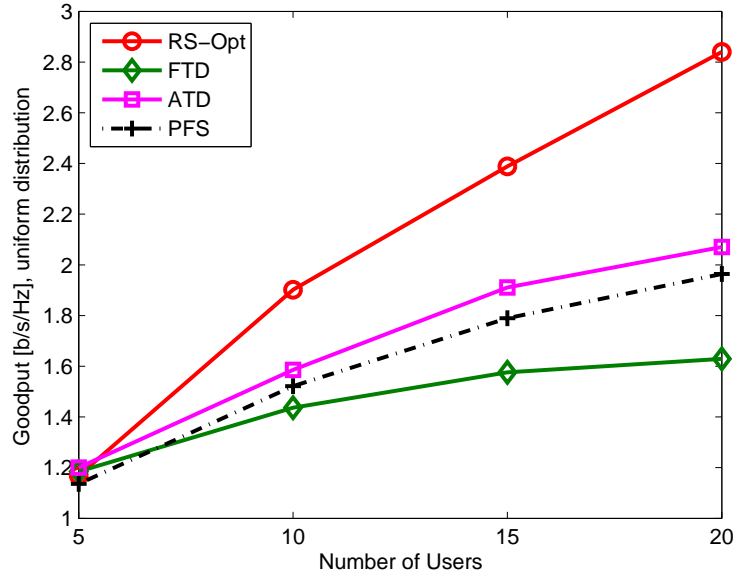


Figure 6.9: Cell goodput in [b/s/Hz] for proposed and reference algorithms, with uniform user distribution, case of single relay

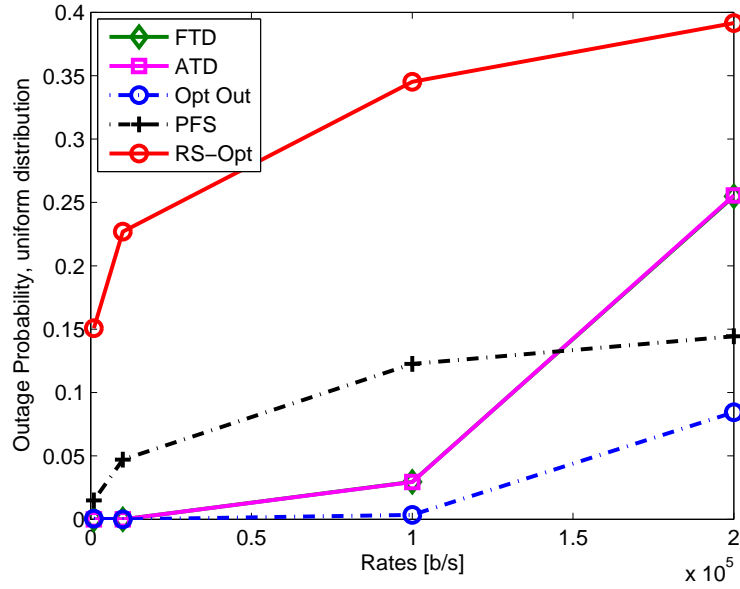


Figure 6.10: Outage Probability for proposed and reference algorithms, with uniform user distribution, case of single relay

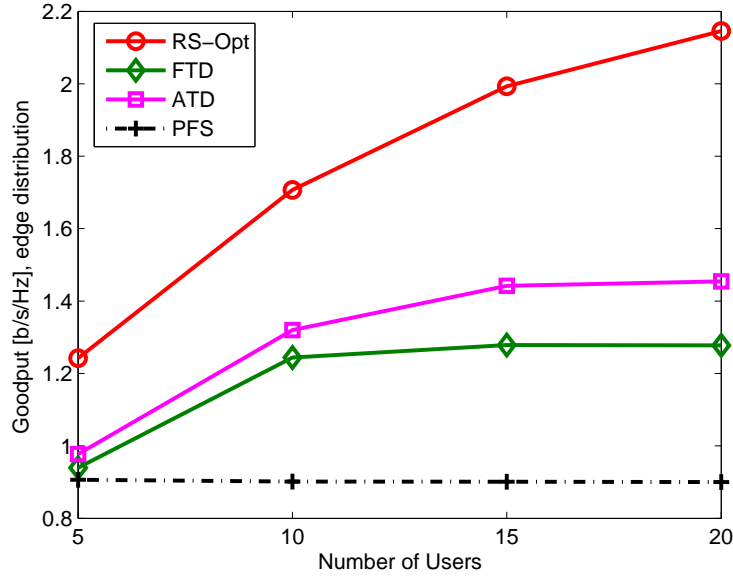


Figure 6.11: Cell goodput in [b/s/Hz] for proposed and reference algorithms, with user distribution towards the cell edge, case of single relay

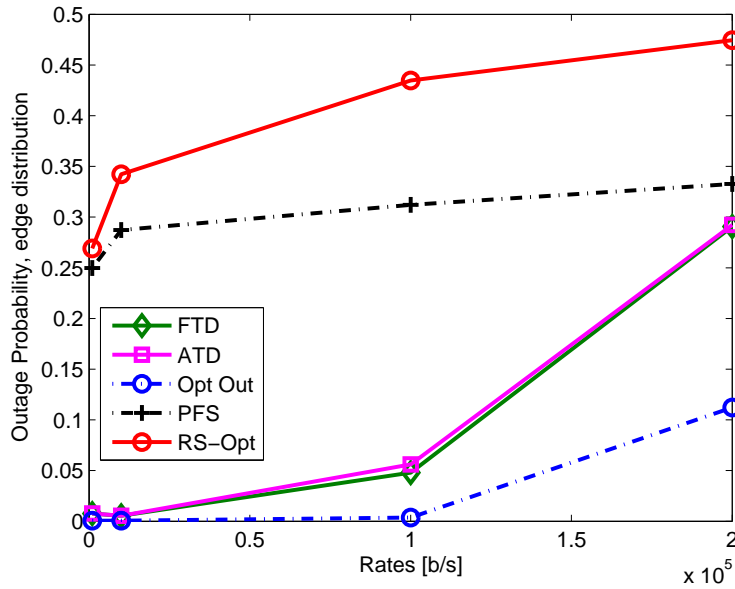


Figure 6.12: Outage Probability for proposed and reference algorithms, with user distribution towards the cell edge, case of single relay

the cell goodput compared to *Conventional MC-PFS* algorithm. The outage performance for FTD and ATD algorithms become also lower than *Conventional MC-PFS* for all values of R . Compared to the uniform distribution, the performance of the *Conventional MC-PFS* algorithm becomes very poor, as shown in Fig. 6.12. Note that since *Optimized Outage* is infeasible, the gap of the proposed algorithms with a practical optimal algorithm would be even smaller. Hence, our algorithms can increase the coverage by increasing the number of users satisfying their rate requirement R . One interesting point is that, when comparing the performance metrics obtained by uniform distribution and edge distribution, the performance in the latter case drops, for all algorithms including the upper and lower bounds. This is due to the fact that the direct link quality of all users is lower in average, since they are farther from the BS. Therefore, the overall throughput and outage deteriorate since there are less users attached to the BS with high link quality, and more users need to be supported by the RS.

From the figures, it can be observed that by adapting the time division of T_{BS} and T_{RS} depending on the queue status at the BS and RS, ATD algorithm significantly improves the throughput while keeping the same outage behavior as the fixed allocation of FTD algorithm. This improvement comes at the price of a higher computational complexity, but which is still much reduced compared to a full search over all the possible time divisions, as discussed in section 6.3.2. On the other hand, FTD achieves good throughput and outage performance, since it outperforms the *Conventional MC-PFS* algorithm with lower complexity and required amount of information. Thus, FTD is well suited for practical use.

Next, the effectiveness of the relay requesting scheme is evaluated and compared to the *All Forward* algorithm presented in section 6.3.3, for uniform user distribution. Figs. 6.13 and 6.14 show that there is a tremendous gain in throughput and in outage with the FTD algorithm, respectively. From Fig. 6.13, it can be seen that as the number of users grows, more and more users are relayed, so that less and less direct users can be allocated in the BS-subframe, which contributes to the overall downfall of the throughput. This confirms the fact that the BS-subframe is over flooded by the relayed users packets, so that most of the subchannels are allocated to the BS-RS link in detriment of the direct users, resulting in this poor performance. By letting the relay choose and request the relayed users' packets, the subchannel allocation between the BS-RS link and the direct users is optimized, which enables a huge performance gain.

6.7.4 Multiple RS Case

With the frequency reuse carried out in schemes such as *MRP* and *MRPA*, some subchannels are used at the same time, so the opposite relays transmitting in parallel interfere with each other. The effect of this interference is taken into account

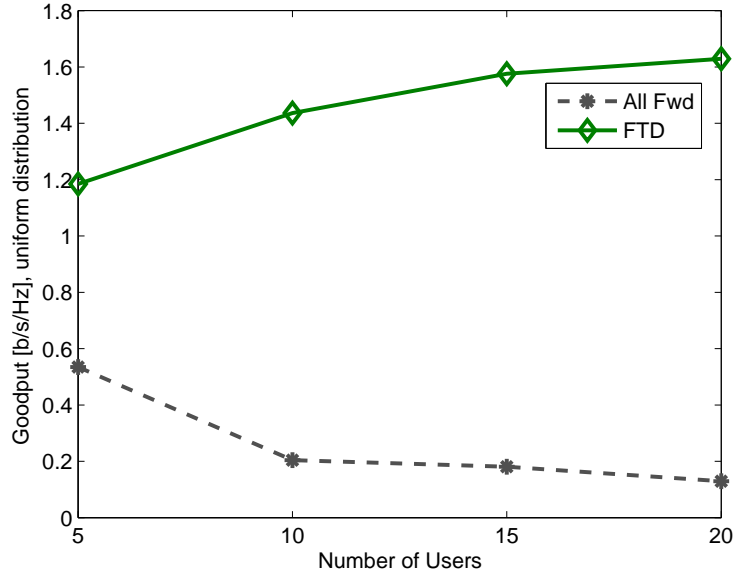


Figure 6.13: Cell goodput in [b/s/Hz] for FTD and All-Forward algorithms, with uniform distribution

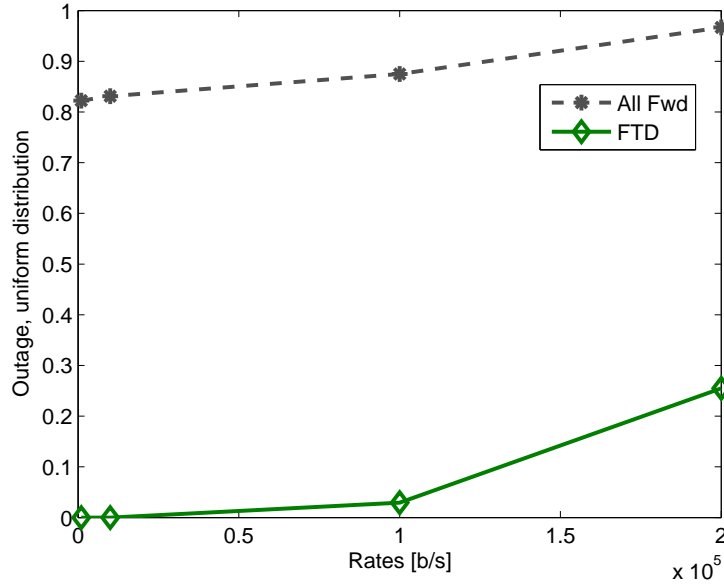


Figure 6.14: Outage Probability for FTD and All-Forward algorithms, with uniform distribution

in the simulations. Namely, if user k attached to RS_1 was scheduled on a subchannel n , then the interference is equal to the signal power of RS_4 to user k , on subchannel n . For the sake of simplicity, the interference is assumed to be an additive Gaussian noise, in which case the SINR of user k on subchannel n denoted $SINR_{k,n}$ is

$$SINR_{k,n} = \frac{p_{k,n} \times |h_{k,n}^{RS_1}|^2}{p_{k,n} \times |h_{k,n}^{RS_4}|^2 + \sigma^2}. \quad (6.15)$$

By denoting $SNR_{k,n}$ the Signal-to-Noise-Ratio (SNR) of user k on subchannel n

$$SNR_{k,n} = \frac{p_n \times |h_{k,n}^{RS_1}|^2}{\sigma^2}, \quad (6.16)$$

we obtain:

$$SINR_{k,n} = \frac{SNR_{k,n}}{1 + SNR_{k,n} \times \frac{|h_{k,n}^{RS_4}|^2}{|h_{k,n}^{RS_1}|^2}}. \quad (6.17)$$

While the subchannel allocation is made based on the SNR values, the SINR values $SINR_{k,n}$ are used to determine the BER values, and thus the achieved throughput.

1. **Results for $I = 6$ relays:** Figs. 6.15 and 6.16 show the goodput and outage performance with uniform user distribution, respectively. It can be first observed from Fig. 6.15 that the goodput achieved by the proposed algorithms is lower than achieved by *RS-Max Full*, as in the single relay case, since *RS-Max Full* is designed to maximize the throughput only, while the proposed ones are based on PFS. At the same time, the outage probability of *RS-Max Full* is considerably higher than compared with the other algorithms. Comparing the proposed algorithms, *MRPA* which uses the activation mechanism and the static allocation scheme *MRP* achieve a similar goodput, while a slight improvement is observed with *MRPA* for a small number of users. That is, if two RS of a same pair are not used, their initial subframe is reallocated among the remaining pairs instead of being wasted. Both *MRP* and *MRPA* have a lower goodput than *Conventional MC-PFS*, but they outperform it in terms of outage, for all values of the required rates R , as shown in Fig. 6.16. Furthermore, with the per-frame RS activation made by *MRAA*, the goodput can be significantly increased as Fig. 6.15 shows, at the cost of outage. Still, *MRAA* outperforms *Conventional MC-PFS* for both metrics, even with the uniform distribution, and achieves the best outage probability for low required rates. The throughput achieved by *Optimized Outage* for all number of users is high since it assumes a full buffer and ideal assumptions.

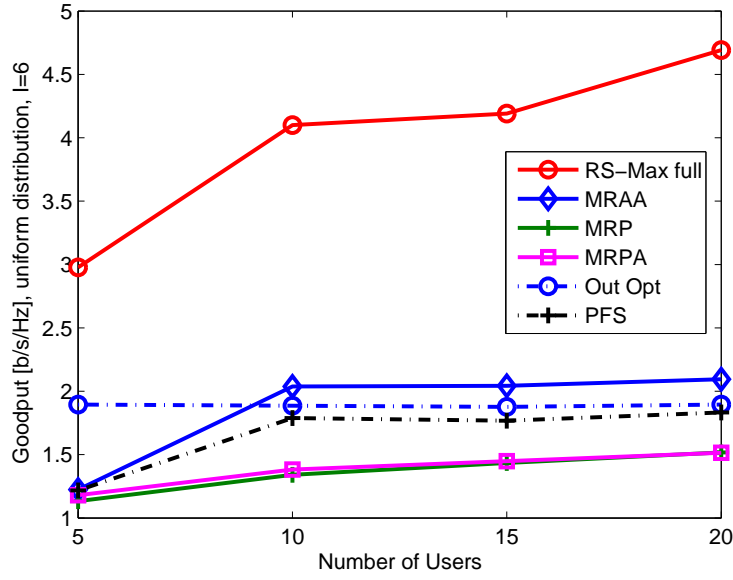


Figure 6.15: Cell goodput in [b/s/Hz] for proposed and reference algorithms, with uniform user distribution

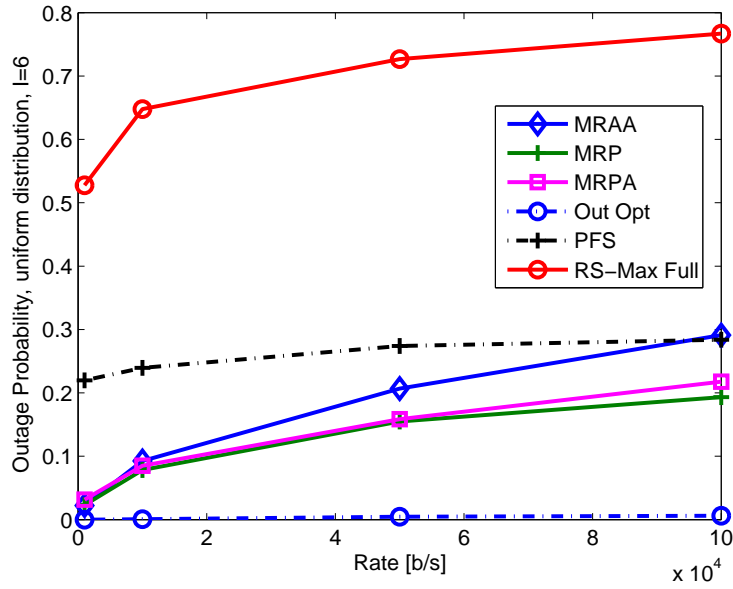


Figure 6.16: Outage Probability for proposed and reference algorithms, with 6 relays, with uniform user distribution

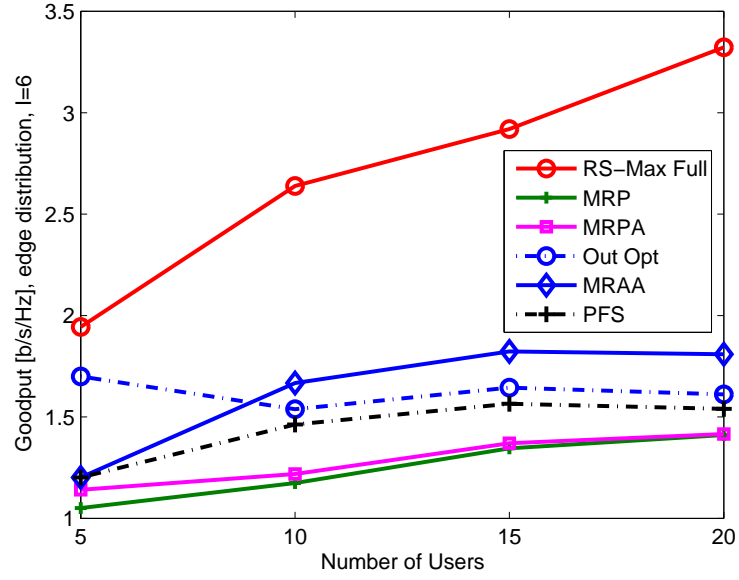


Figure 6.17: Cell goodput in [b/s/Hz] for proposed and reference algorithms, with 6 relays and users distributed towards cell edge

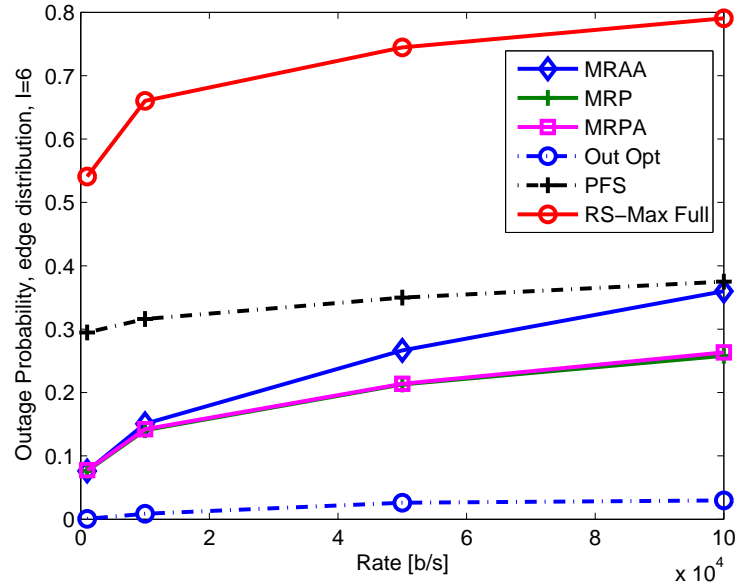


Figure 6.18: Outage Probability for proposed and reference algorithms, with 6 relays and users distributed towards cell edge

When the users are distributed towards the cell edge, it can be first noticed from Figs. 6.17 and 6.18, that the overall performance of all the algorithms is decreased for both metrics, which is due to the same reason as for the single relay case, e. g., the general degradation of the direct user links. In particular, Fig. 6.18 shows that the outage performance of *Conventional MC-PFS* is considerably degraded compared to the uniform user distribution. However, the outage performance of the proposed algorithms is only slightly higher than in the uniform distribution case, while their relative gain from the outage of *Optimized Outage* remains similar. Thus, all the proposed algorithms provide a significant outage reduction compared to *Conventional MC-PFS*, while *MRAA* outperforms *Conventional MC-PFS* for both goodput and outage, for different types of user distributions. In other words, the proposed algorithms are robust with respect to varying user distributions.

Finally, for both user distributions, the comparison between the proposed algorithms shows that the static allocation by *MRP* algorithm is quite inefficient, regarding the goodput performance, due to the empty RS_i -subframes. This validates the fact that by introducing the time adaptation with a little bit higher complexity, the performance can be much improved. Thus, *MRAA* achieves the best goodput/outage trade-off, even surpassing the schemes with frequency reuse, namely *MRP* and *MRPA*.

2. **Results for varying I :** In this part, goodput and outage probabilities are plotted in function of the number of relays with $K = 20$ users. For the outage probability, the target rate is fixed to 100 kbps. In these evaluations, the amount of resource used for the preambles and the user mapping information vary with the number of relays. The impact of these control fields is taken into account in the overall throughput. Simply, there is one preamble per relay which occupies one OFDM symbol, $n_{pre} = 1$. In case of frequency reuse by opposite relays, their corresponding preambles also reuse the frequencies. The DL mapping information consists of the ID of the allocated user (or relay for the BS-RS link), for each subchannel. Denoting n_{MAP} the number of OFDM symbols used for the DL mapping, the goodput is determined as

$$\gamma = \tau \times \frac{n_{data}}{n_{data} + n_{OH} + (N_{RS} + 1) \times (n_{pre} + n_{MAP})}, \quad (6.18)$$

where N_{RS} is the number of non-overlapped, activated RS-Subframes.

Figs. 6.19 and 6.20 show the system goodput and outage performance for uniform user distribution, respectively. The goodput and outage of *Conventional MC-PFS* algorithm remain constant since this algorithm is indepen-

dent from the number of relays in the cell. It is observed from Fig. 6.19 that *RS-Max full* algorithm achieves the best goodput when $I = 6$ relays. The goodput decrease of *RS-Max Full* algorithm at $I = 8, 12$ relays is due to the higher amount of signalling which comes with a higher number of relays.

It is observed that for the proposed algorithms and *Optimized Outage*, the performance of both metrics do not necessarily improve as the number of relays increase. For *Optimized Outage*, the outage at $I = 6$ is already so low that increasing the number of relays does not provide noticeable improvement. Among the proposed algorithms, *MRP* and *MRPA* achieve the best outage performance for all number of relays and outperform *Conventional MC-PFS*, but their best goodput occurs for $I = 8$ relays. This is because increasing the number of relays to $I = 12$ reduces the time allocated to each RS_i subframe. The worst performance of *MRP* and *MRPA* occurs for $I = 3$, which can be explained by the fact that there is no frequency reuse. The improvement in goodput and outage for $I \geq 6$ stresses the effect of the frequency reuse. However, *MRAA* outperforms *Conventional MC-PFS* in terms of both goodput and outage, for all number of relays. It can be observed from Fig. 6.19 that the goodput of *MRAA* increases with the number of relays, even though the amount of signalling increases as shown in (6.18), which indicates the robustness of the proposed scheme to increasing signalling overhead. At the same time, the outage probability of *MRAA* increases as the number of relays grow for $I \geq 6$, since this algorithm is designed to allocate in each frame, the relays which increase the overall throughput. As the number of relays grow, the number of non-allocated relays increase, causing the outage to rise. Still, the outage performance of *MRAA* at $I = 12$ is considerably lower than the one of *Conventional MC-PFS*, being located half-way between the lower bound outage and the outage of *Conventional MC-PFS*. Besides, note that *Optimized Outage* is an infeasible scheme so that the difference with the actual optimal scheme is smaller.

When users are distributed towards the edge, the overall performance of all the algorithms is again generally degraded for all number of relays, as observed in Figs. 6.21 and 6.22. For all number of relays, while the outage lower bound of *Conventional MC-PFS* is largely degraded, the outage of the proposed algorithms is only slightly lower than with uniform user distribution, pointing out their robustness regarding varying user distributions. Comparing the proposed algorithms with *Conventional MC-PFS*, the same conclusions can be made as in the uniform user distribution case. The gain in goodput of *MRAA* compared to *Conventional MC-PFS* is even higher than with the uniform distribution, and now *MRP* and *MRPA* also outper-

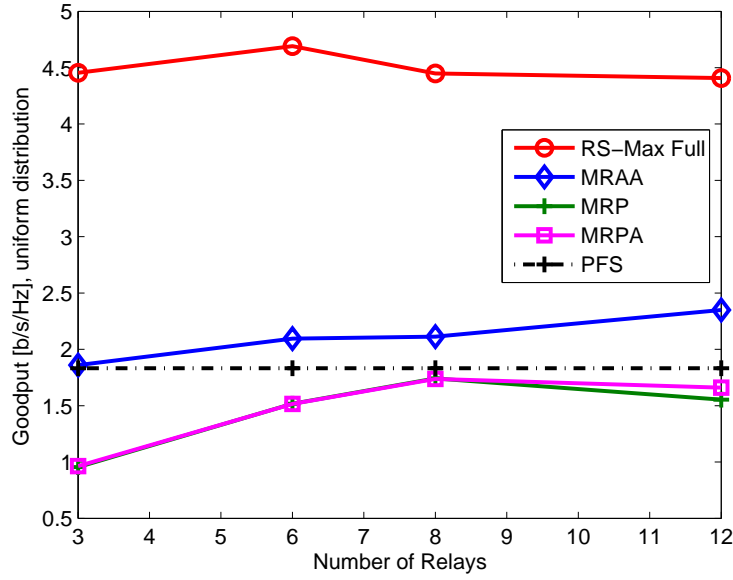


Figure 6.19: Cell goodput in [b/s/Hz] for proposed and reference algorithms with varying number of relays, uniform user distribution

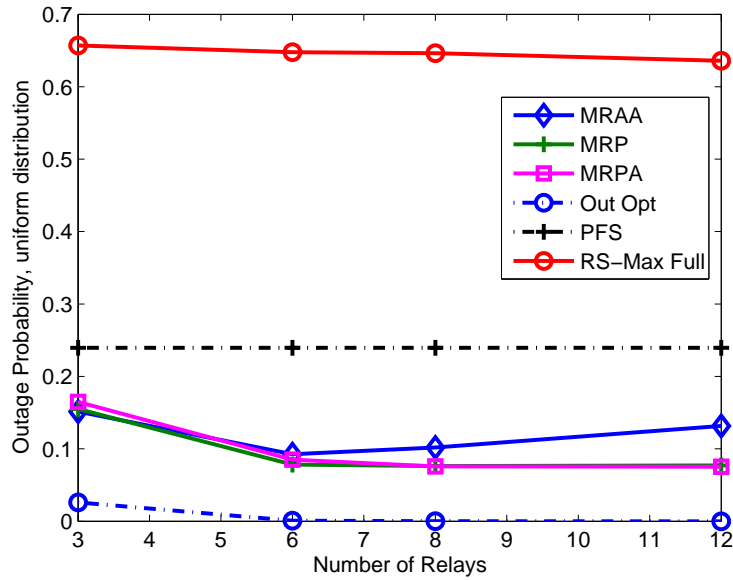


Figure 6.20: Outage Probability for proposed and reference algorithms with varying number of relays, uniform user distribution

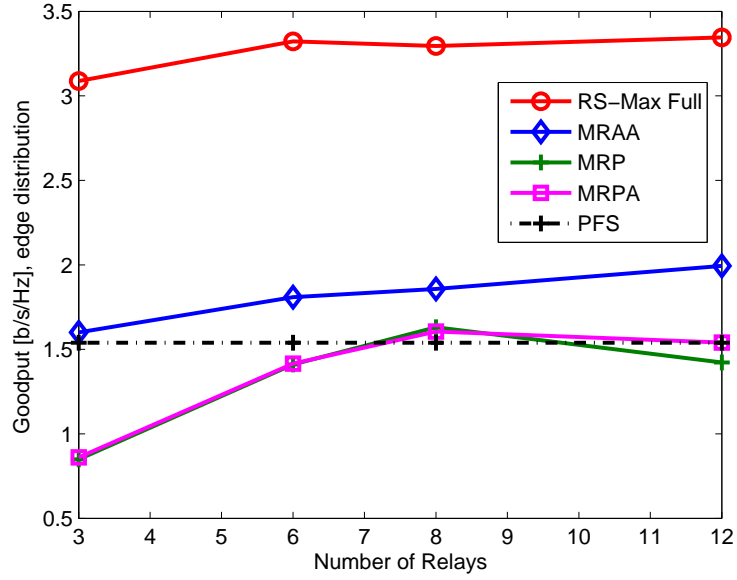


Figure 6.21: Cell goodput in [b/s/Hz] for proposed and reference algorithms with varying number of relays, users distributed towards cell edge

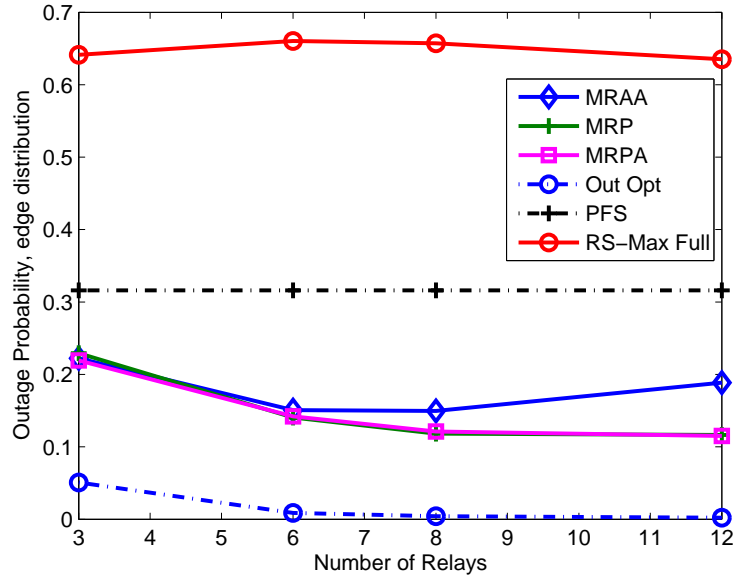


Figure 6.22: Outage Probability for proposed and reference algorithms with varying number of relays, users distributed towards cell edge

form *Conventional MC-PFS* for $I = 6$, while achieving the best outage for all I . However, the goodput of *MRP* suddenly degrades when going from $I = 8$ to $I = 12$ relays when users are distributed towards the edge. This can be explained by the fact that the number of activated relays increase as more users are located in the cell outskirts, which increases the amount of signalling overhead and decreases the time allocated to each RS_i -subframe.

The simulations with different number of relays in the cell showed the efficiency of the proposed algorithms, which are robust to varying number of relays and to different user distributions. The schemes *MRP* and *MRPA* achieve the best outage performance, while *MRAA* offers significant goodput increase, at the expense of a slightly higher but reasonable complexity. Thus, *MRAA* achieves the best performance with an excellent trade-off between goodput and outage, for different number of users, relays, and is robust to different user distributions. Under the considered assumptions, the optimal number of relays seem to differ depending on the algorithm. *MRP* and *MRPA* achieve their best performance for $I = 8$, for both user distributions. For *MRAA*, $I = 8$ also achieves the best compromise between goodput and outage, but $I = 12$ offers the best goodput level while considerably reducing the outage probability compared to the conventional algorithm. Interestingly, continuously increasing the number of relays may not be the best solution even when there are more users in the cell edge, since, for example, the goodput of *MRP* and *MRPA* is affected by the increasing amount of signalling overhead as more and more relays are activated and less resource is allocated to each RS_i -subframe. As the number of relays increase, the goodput of *MRAA* improves but not the outage since more and more RS_i -subframes are not allocated.

6.8 Summary

In this chapter, we have investigated the problem of resource allocation for a relay-aided cellular system based on OFDMA. Algorithms for resource allocation have been proposed, for single RS and multiple RS cases. The main specificity of these algorithms is that the RS makes its own initial allocation to minimize the outage, and thereafter the BS optimizes the final allocation in order to improve the overall throughput. This allows a tremendous decrease of the required signalling. In the single RS case, two algorithms have been designed: one operating with low complexity by keeping a fixed time division and the second, with a higher complexity as the time is adapted. In the multiple RS case, the concept of RS activation has been proposed, where the frame structure is adapted depending on the active RS. For different number of relays in the cell, and different types of user dis-

tributions, the simulation results show that the proposed schemes achieve a good trade-off between outage and throughput compared to reference algorithms, while minimizing the computational complexity and required amount of CSI. Depending on the number of relays, the advantages of each algorithm were discussed. Without a dynamic selection of the useful relays, the system hardly benefits from the static scheme when the number of relays increase, even with frequency reuse for the static scheme. This indicates that adaptive schemes such as the proposed ones are necessary in order provide an overall performance enhancement.

Algorithms providing a trade-off between competing performance measures, such as throughput and outage, are very useful. Moreover, these are low-complexity algorithms which make them even more suited for practical use.

Chapter 7

Amplify-and-Forward Cooperative Diversity Schemes for Multi-Carrier Systems

7.1 Introduction and Motivation

Cooperative diversity has been subject of a number of works, as discussed in section 2.9, which mostly focus on the physical layer issues in a SC system. However, there are still unresolved problems when taking the higher layer point of view, especially in a MC system. When several relays are potential candidates for cooperation, which of them should participate? and how the radio resource should be shared among them? [57] addresses this issue for a SC, AF system by proposing two types of allocation, namely the All-Participate-AF (AP-AF) scheme, shown in Fig. 7.1, where the available channel resource, in either time or frequency, is split between the I relays and the source, and the Selection-AF (S-AF) scheme, shown in Fig. 7.2, where the available resource is equally divided among the source and the best relay among I (R_2 in this example). It is shown in [57] that both schemes achieve the full diversity order of $(I + 1)$, while S-AF improves considerably the outage probability, defined as the probability that a predefined target rate is not attained. This is because the throughput performance of the AP-AF scheme is decreased since the resource is orthogonally partitioned.

However, those studies do not consider a MC setting, where parallel channel resources are available at the same time. Since MC transmission is likely to become a key element in the future wireless communication system, it is of greatest interest to examine the potential of cooperative diversity techniques in a MC system. The results obtained in [57] for the SC case can not be simply extended to the MC case, as much more involved analysis are required when the additional

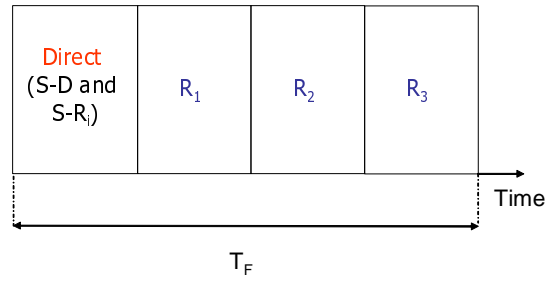


Figure 7.1: Frame Structure of the All-Participate-AF (AP-AF) scheme

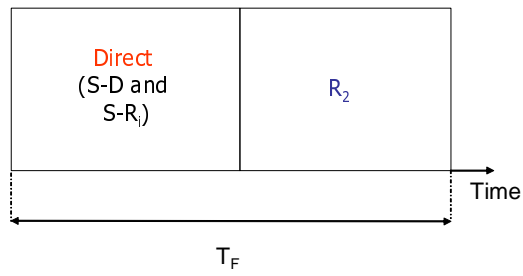


Figure 7.2: Frame Structure of the Selection-AF (S-AF) scheme

degree of freedom in frequency comes into play. In [41], the system considered is equivalent to the Frequency Division Multiple Access (FDMA) case, where one subchannel is allocated to each transmitter, which allows to limit the analysis to one subchannel, as opposed to the MC case. Moreover, the theoretical analysis is limited to the DF protocol, where the S–R and R–D channels can be modeled independently. In this work, we will focus on the AF protocol, hence the results in [41] can not be straightforwardly translated. Moreover, the main feature of the AF scheme is the simplicity of the relays, which should only amplify and forward the data sent by the source, without performing any other complex operation. Such an assumption makes restrictions on the space of possible transmission strategies that can be used. Thus, it is assumed that the relays do not provide any CSI feedback to the source. Moreover, the relays do not perform any allocation or reordering of the subcarriers. Under such conditions, the whole decision process is left to the source and the destination. Thus, CSI feedback is considered from the destination to the source. This is another major concern in the MC case, since the source needs the knowledge of all the channels (relayed and direct) in all subcarriers in order to optimize capacity or system outage. This becomes rapidly prohibitive as the number of subcarriers and relays grow. Thus, the cooperative diversity schemes should be designed taking into account those signalling issues, while providing a good performance.

In this chapter, after introducing the system model, we present the proposed allocation schemes for MC systems. This is followed by a theoretical analysis, where the bounding expressions of the system outage are derived for all the schemes. Computer simulations show the validity of the theoretical analysis, and also when considering more practical assumptions such as real OFDM channels.

7.2 System Model

The system considered in this work is depicted in Fig. 7.3. The source S communicates with the destination D via I relays R_i , $i \in \{1, \dots, I\}$ where $I = 3$ in this example. Transmissions from the relays can be organized in different ways, which will be explained in the following sections. Depending on the cooperative diversity scheme, the transmission occurs in 2 to $I + 1$ phases. In the first phase, the source transmits to the destination and relays. In the subsequent phases, depending on the scheme, i relays among I transmit to the destination. In phase 1 of each frame, the source S transmits an OFDM symbol of N subcarriers with power P_s . $P_{s,n}$ is the amount of power allocated to subcarrier n , so that $P_s = \sum_{n=1}^N P_{s,n}$.

Since consider diversity schemes are considered, the transmitted bits are multiplexed over all the subcarriers which are applied the same modulation level, corresponding to their average SNR level. It is also possible to perform per subcarrier

bit loading where the modulation level is adapted per subcarrier, but as mentioned in [41], this leads to a more complex system than the diversity setting, and in addition, the source needs to know the SNR values in each subcarrier, resulting into prohibitively high amounts of CSI feedback. Therefore, the bit loading schemes are beyond the scope of this chapter.

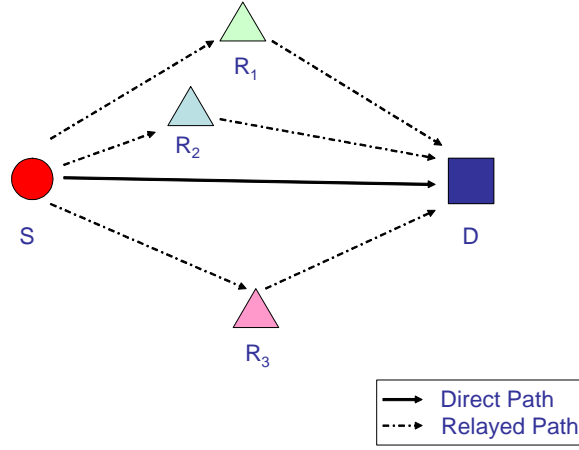


Figure 7.3: System Model

In phase 1, by assuming equal power distribution, the received signals at the destination D and relay R_i in the n -th subcarrier, $n \in \{1, \dots, N\}$, can be written as

$$\begin{aligned} y_{s,d,n} &= \sqrt{P_{s,n}} h_{s,d,n} x_n + w_{s,d,n} \\ y_{s,i,n} &= \sqrt{P_{s,n}} h_{s,i,n} x_n + w_{s,i,n} \end{aligned} \quad (7.1)$$

$h_{s,d,n}$, $h_{s,i,n}$ and $h_{i,d,n}$ are the channel fading gains of (S,D) , (S,R_i) and (R_i,D) on subcarrier n , respectively, modeled as a circular symmetric complex Gaussian random variable with mean zero and variance one, namely $h_{s,d,n}, h_{s,i,n}, h_{i,d,n} \sim CN(0, 1)$. $w_{s,d,n} \sim CN(0, N_{s,d})$ and $w_{s,i,n} \sim CN(0, N_{s,i})$ are the Additive White Gaussian Noise (AWGN), and x_n denotes the symbol on subcarrier n .

In phases 2 to $I + 1$, each relay amplifies $y_{s,i,n}$ and transmits it to D with transmission power $P_{i,n}$. The received signal at destination D from relay R_i in subcarrier

n is given by

$$\begin{aligned} y_{i,d,n} &= \sqrt{P_{i,n}} h_{i,d,n} \frac{y_{s,i,n}}{\sqrt{E[|y_{s,i}|^2]}} + w_{i,d,n} \\ &= \sqrt{\frac{P_{s,n} P_{i,n}}{P_{s,n} |h_{s,i,n}|^2 + N_{s,i}}} h_{s,i,n} h_{i,d,n} x_n + \tilde{w}_{i,d,n}, \end{aligned} \quad (7.2)$$

where the operator $E[\cdot]$ denotes the ensemble average, $w_{i,d,n} \sim CN(0, N_{i,d})$ denotes the AWGN in the (R_i, D) channel in subcarrier n , and

$$\tilde{w}_{i,d,n} = \sqrt{\frac{P_{i,n}}{P_{s,n} |h_{s,i,n}|^2 + N_{s,i}}} h_{i,d,n} w_{s,i,n} + w_{i,d,n} \quad (7.3)$$

is the equivalent noise, with zero mean and variance $\tilde{N}_{i,d}$. The received signal in subcarrier n , from all the relays can be written in vector form as in [57]

$$\mathbf{y}_{d,n} = \mathbf{h}_n x_n + \mathbf{w}_n, \quad (7.4)$$

with $\mathbf{w}_n \sim CN(\mathbf{0}, \mathbf{I})$ and where

$$\mathbf{y}_{d,n} = \begin{bmatrix} \frac{y_{s,d,n}}{\sqrt{N_{s,d}}} & \frac{y_{1,d,n}}{\sqrt{\tilde{N}_{1,d}}} & \cdots & \frac{y_{I,d,n}}{\sqrt{\tilde{N}_{I,d}}} \end{bmatrix}^T \quad (7.5)$$

$$\mathbf{h}_n = \begin{bmatrix} \frac{\sqrt{P_{s,n}} h_{s,d,n}}{N_{s,d}} & \frac{1}{\sqrt{\tilde{N}_{1,d}}} \sqrt{\frac{P_{s,n} P_{1,n}}{P_{s,n} |h_{s,1,n}|^2 + N_{s,1}}} h_{1,d,n} h_{s,1,n} \\ \cdots & \frac{1}{\sqrt{\tilde{N}_{I,d}}} \sqrt{\frac{P_{s,n} P_{I,n}}{P_{s,n} |h_{s,I,n}|^2 + N_{s,I}}} h_{I,d,n} h_{s,I,n} \end{bmatrix}^T \quad (7.6)$$

By assuming ideal Maximal Ratio Combining (MRC) at the destination, the SNR available at the combiner output is the sum of the instantaneous SNR of all subcarriers and relays.

7.3 Proposed Cooperative Diversity Schemes

7.3.1 All-Participate All Subcarrier Scheme (APN) and All-Participate Rate Splitting Scheme (AP-RS)

Both schemes, APN and AP-RS share the same frame structure depicted in Fig. 7.4, for $I = 3$. The time frame, with duration T_F , is equally divided between the source and all relays. In the first phase, referred to as direct transmission phase, the

source S transmits to the destination (S,D) and to the relays (S,R_i). In the subsequent phases, the I relays transmit sequentially. The source and relays use all the subcarriers within their allocated frame portion. When all the relays participate, there are two different ways to multiplex the data over the subcarriers:

1. Spreading the data accross all the subcarriers before sending it. This is referred as the APN scheme. This scheme will be in outage if the total rate resulting from the sum of the SNR of all subcarriers is smaller than the target rate. The APN scheme works as the AP-AF scheme presented in [57], but adapted to the MC case.
2. Splitting the total data into N equal streams which are transmitted from each subcarrier at the same rate, equal to the target rate R divided by N . This is referred to as the AP-RS scheme. Since all the subcarriers have to carry a data rate larger than R/N , this scheme will be in outage if any of the subcarriers has a rate smaller than R/N .

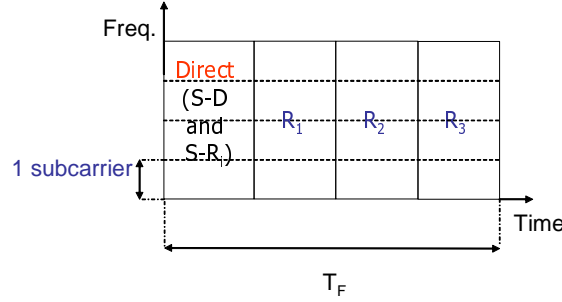


Figure 7.4: Frame Structure for the APN Scheme

7.3.2 Average Best Relay Selection Scheme (AvgBRS)

This second scheme is the extension to the MC case of the S-AF scheme introduced in [57]. As depicted in Fig. 7.5, the relay with the best SNR averaged over all the subcarriers is chosen as the transmitting relay, in this case it is R_2 . The frame is divided in two, resulting in a two-phased transmission.

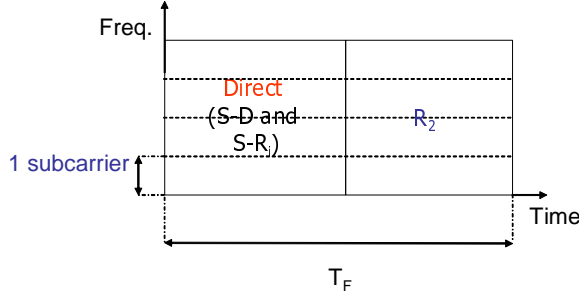


Figure 7.5: Frame Structure for the AvgBRS Scheme

7.3.3 Per-Subcarrier Best Relay Selection Scheme (NBRS)

In this third scheme shown in Fig. 7.6, the transmission occurs also in two phases: in the first phase, the source transmits, and in the second, the relays with the best instantaneous end-to-end SNR are assigned to the corresponding subcarrier. Hence, this scheme may not be practical since it requires a higher complexity and the knowledge of the CSI for all relays in all subcarriers, which must be sent to the source via an uplink feedback channel in order to perform the allocation. However, NBRS gives a better performance compared to APN and AvgBRS since the allocation is adapted per-subcarrier. In this scheme, different relays are allocated to each subcarrier but this allocation remains constant for the two transmission phases. It could be also possible to have different per subcarrier relay allocations in each phase, by means of subcarrier reordering at the relays. But, as mentioned in the introduction, since we focus on the AF protocol, it will not be considered. That is, if the subcarrier is changed at the relay, the complexity of this operation in AF mode becomes comparable to that required by DF mode, thereby hindering the main advantage of the AF protocol which is to minimize the processing complexity by only amplifying the signal instead of decoding it. However, such a scheme could be analyzed for getting the optimal performance but it is not considered here, as it seems far from practical.

7.3.4 Random Relay Selection Scheme (RRS)

This scheme is designed to provide a reference to the performance of the schemes presented above. We have a two-phased transmission, where the source transmits

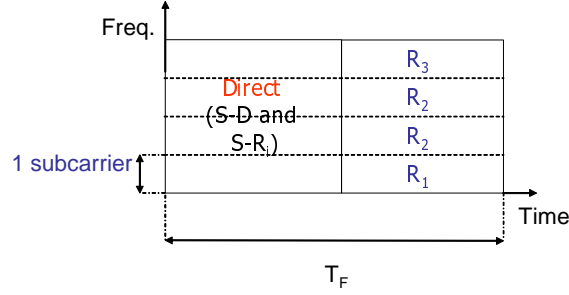


Figure 7.6: Frame Structure for the NBRs Scheme

during the first half of the frame, and then a randomly chosen relay transmits in the second half, in all the subcarriers.

7.4 Preliminary Analysis

First let us consider the following theorem which will be used in the analysis of the outage probability in the following section. This result is a generalization of the theorem in [22], which can be applied when there are multiple carriers and/or relays.

Theorem 1. Let $r_\delta(N) = \sum_{n=1}^N \delta f(v_n/\delta, w_n/\delta)$ with δ positive, where $f(x, y) = (xy)/(x + y + 1)$ and v_n and w_n are independent exponential random variables with parameters $\lambda_n, \mu_n \forall n$, respectively. Let $h(\delta) > 0$ be continuous with $h(\delta) \rightarrow 0$ and $\delta/h(\delta) \rightarrow k < \infty$ as $\delta \rightarrow 0$. Then the probability $Pr[r_\delta(N) < \delta]$ satisfies

$$\liminf_{\delta \rightarrow 0} \frac{1}{h^N(\delta)} Pr[r_\delta(N) < h(\delta)] \geq \prod_{n=1}^N \left(\frac{\lambda_n + \mu_n}{2} \right)$$

$$\limsup_{\delta \rightarrow 0} \frac{1}{h^N(\delta)} Pr[r_\delta(N) < h(\delta)] \leq \prod_{n=1}^N (\lambda_n + \mu_n).$$

Proof. See Appendix B. □

For the SC case, where $r_\delta(1) = \delta f(v_1/\delta, w_1/\delta)$ and λ_1, μ_1 are the parameters of the exponential random variables v_1 and w_1 , it was shown in [22] that

$$\lim_{\delta \rightarrow 0} \frac{1}{h(\delta)} Pr[r_\delta(1) < h(\delta)] = \lambda_1 + \mu_1. \quad (7.7)$$

7.5 Theoretical Analysis of Outage Probability

In this section, the outage probability of the proposed schemes is theoretically analyzed. Since it is very difficult to obtain the exact closed-form expressions, the analysis is conducted by deriving the upper and lower bounds of the outage in the high SNR region, using the results presented in section 7.4.

7.5.1 All-Participate All Subcarrier Scheme (APN)

The capacity achieved by APN can be written as

$$\begin{aligned} C_{APN} &= \frac{1}{I+1} \log_2 \left(1 + \sum_{n=1}^N \mathbf{h}_n^* \mathbf{h}_n \right) \\ &= \frac{1}{I+1} \log_2 \left(1 + \sum_{n=1}^N \frac{P_{s,n} |h_{s,d,n}|^2}{N_{s,d}} + \sum_{i=1}^I \sum_{n=1}^N \frac{P_{s,n} P_{i,n} |h_{i,d,n}|^2 |h_{s,i,n}|^2}{\tilde{N}_{i,d} (P_{s,n} |h_{s,i,n}|^2 + N_{s,i})} \right), \end{aligned} \quad (7.8)$$

where $[\cdot]^*$ denotes the Hermitian transpose. By defining $N_{s,d} = N_0$, $N_{s,i} = k_{s,i} N_0$, $N_{i,d} = k_{i,d} N_0$ and $\alpha_{0,n} = P_{s,n} |h_{s,d,n}|^2$, $\alpha_{i,n} = P_{s,n} |h_{s,i,n}|^2 / k_{s,i}$ and $\beta_{i,n} = P_{s,n} |h_{i,d,n}|^2 / k_{i,d}$, posing $\gamma = 1/N_0$ as in [57] for the system SNR, we can rewrite (7.8) as

$$C_{APN} = \frac{1}{I+1} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right). \quad (7.9)$$

The outage probability, which is the probability that the capacity C_{APN} falls below a predetermined rate R , becomes

$$\begin{aligned} P_{out}^{APN} &= P \left[\frac{1}{I+1} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right) < R \right] \\ &= P \left[\sum_{n=1}^N \alpha_{0,n} + \sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \tilde{\delta} \right], \end{aligned} \quad (7.10)$$

where $\tilde{\delta}$ is defined as

$$\tilde{\delta} = \frac{2^{(I+1)R} - 1}{\gamma}. \quad (7.11)$$

Since we assume Rayleigh fading channels $h_{s,d,n}, h_{s,i,n}, h_{i,d,n} \sim CN(0, 1)$, the weighted amplitude squares $\alpha_{0,n}$, $\alpha_{i,n}$ and $\beta_{i,n}$ are exponentially distributed with parameters $\lambda_{0,n}=\lambda_0$, $\lambda_{i,n}$ and $\mu_{i,n}$, respectively, defined as

$$\begin{aligned}\lambda_0 &= \frac{1}{P_{s,n}E[|h_{s,d,n}|^2]} = \frac{1}{P_{s,n}} \\ \lambda_{i,n} &= \frac{1}{P_{s,n}/k_{s,i}E[|h_{s,i,n}|^2]} = \frac{1}{P_{s,n}/k_{s,i}} \\ \mu_{i,n} &= \frac{1}{P_{s,n}/k_{i,d}E[|h_{i,d,n}|^2]} = \frac{1}{P_{s,n}/k_{i,d}}.\end{aligned}\quad (7.12)$$

As a sum of N independent random exponential variables with parameter λ_0 , the probability density function of $x=\sum_{n=1}^N \alpha_{0,n}$ follows a chi-square distribution and is equal to

$$p_x(x) = \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x} x^{N-1}, \quad (7.13)$$

by applying successively the convolution formula for a sum of independent random variables [58].

Thus,

$$P_{out}^{APN} = \int_0^{\tilde{\delta}} P\left[\sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \tilde{\delta} - x\right] \times \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x} x^{N-1} dx. \quad (7.14)$$

With a change of variable $x'=1 - x/\tilde{\delta}$,

$$\begin{aligned}P_{out}^{APN} &= \int_0^1 P\left[\sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \tilde{\delta} x'\right] \times \tilde{\delta}^N \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x \tilde{\delta}(1-x')}(1-x')^{N-1} dx' \\ &= \tilde{\delta}^{(I+1)N} \int_0^1 \left(\frac{P\left[\sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \tilde{\delta} x'\right]}{(\tilde{\delta} x')^{IN}} \right) \\ &\quad \times \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x \tilde{\delta}(1-x')} x'^{IN} (1-x')^{N-1} dx'\end{aligned}\quad (7.15)$$

We observe

$$\lim_{\gamma \rightarrow \infty} e^{-\lambda_0 x \tilde{\delta}(1-x')} = 1, \quad (7.16)$$

since $\tilde{\delta} \rightarrow 0$ as $\gamma \rightarrow \infty$.

As Theorem 1 stipulates

$$\liminf_{\gamma \rightarrow \infty} \frac{P\left[\sum_{i=1}^I \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \tilde{\delta} x'\right]}{(\tilde{\delta} x')^{IN}} \geq \prod_{i=1}^I \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2}\right), \quad (7.17)$$

we obtain a lower bound L_{APN} for outage probability P_{out}^{APN} for high SNR, as

$$L_{APN} = \tilde{\delta}^{(I+1)N} \frac{\lambda_0^N}{(N-1)!} \prod_{i=1}^I \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) \times \int_0^1 x'^{IN} (1-x')^{N-1} dx' \quad (7.18)$$

Finally, since $\int_0^1 x'^{IN} (1-x')^{N-1} dx' = \frac{(N-1)!(IN)!}{(IN+N)!}$, we obtain

$$L_{APN} = \frac{\lambda_0^N (IN)!}{(IN+N)!} \prod_{i=1}^I \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) \tilde{\delta}^{(I+1)N}. \quad (7.19)$$

We also obtain an upper bound U_{APN} for outage probability P_{out}^{APN} using Theorem 1 as

$$U_{APN} = \frac{\lambda_0^N (IN)!}{(IN+N)!} \prod_{i=1}^I \prod_{n=1}^N (\lambda_{i,n} + \mu_{i,n}) \tilde{\delta}^{(I+1)N}. \quad (7.20)$$

7.5.2 All-Participate Rate Splitting Scheme (AP-RS)

As explained in section 7.3, this scheme is in outage if one of the subcarriers carry a rate smaller than the target rate divided by the number of subcarriers. If r_n is the rate carried by subcarrier n , the outage probability can be written

$$P_{out}^{APRS} = 1 - P\left[\forall n \in [1..N], r_n \geq \frac{R(I+1)}{N}\right], \quad (7.21)$$

since each subcarrier is shared between the all the relays and the source. By assuming the subcarriers independent, P_{out}^{APRS} becomes

$$\begin{aligned} P_{out}^{APRS} &= 1 - \prod_{n=1}^N P\left[r_n \geq \frac{R(I+1)}{N}\right] \\ &= 1 - \prod_{n=1}^N P\left[\log_2 \left(1 + \gamma\alpha_{0,n} + \sum_{i=1}^I \frac{\gamma\alpha_{i,n}\gamma\beta_{i,n}}{1 + \gamma\alpha_{i,n} + \gamma\beta_{i,n}}\right) \geq \frac{R(I+1)}{N}\right] \\ &= 1 - \prod_{n=1}^N P\left[\sum_{i=1}^I \frac{\alpha_{i,n}\gamma\beta_{i,n}}{1 + \gamma\alpha_{i,n} + \gamma\beta_{i,n}} \geq \bar{\delta} - \alpha_{0,n}\right], \end{aligned} \quad (7.22)$$

where $\bar{\delta}$ is defined as

$$\bar{\delta} = \frac{2^{\frac{R(I+1)}{N}} - 1}{\gamma}. \quad (7.23)$$

With the same method as in section 7.5.1, using Theorem 1 we obtain the bounds which can be expressed

$$L_{APRS} = 1 - \prod_{n=1}^N \left[1 - \frac{\lambda_0}{I+1} \prod_{i=1}^I \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) \times \bar{\delta}^{I+1} \right] \quad (7.24)$$

$$U_{APRS} = 1 - \prod_{n=1}^N \left[1 - \frac{\lambda_0}{I+1} \prod_{i=1}^I (\lambda_{i,n} + \mu_{i,n}) \times \bar{\delta}^{I+1} \right]. \quad (7.25)$$

7.5.3 Average Best Relay Selection Scheme (AvgBRS)

The capacity achieved by AvgBRS can be written as

$$C_{AvgBRS} = \frac{1}{2} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \max_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right), \quad (7.26)$$

and the outage probability is written

$$\begin{aligned} P_{out}^{AvgBRS} &= P \left[\frac{1}{2} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \max_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right) < R \right] \\ &= P \left[\sum_{n=1}^N \alpha_{0,n} + \max_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} \right], \end{aligned} \quad (7.27)$$

where $\hat{\delta}$ is defined as

$$\hat{\delta} = \frac{2^{2R} - 1}{\gamma}. \quad (7.28)$$

In the same way as for APN scheme, we derive the bounds for AvgBRS scheme. P_{out}^{AvgBRS} becomes

$$P_{out}^{AvgBRS} = \int_0^{\hat{\delta}} P \left[\max_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} - x \right] \times \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x} x^{N-1} dx. \quad (7.29)$$

After a change of variable $x' = 1 - x/\hat{\delta}$, we can write

$$\begin{aligned} P_{out}^{AvgBRS} &= \hat{\delta}^{(I+1)N} \int_0^1 \prod_{i=1}^I \left(\frac{P \left[\sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} x' \right]}{(\hat{\delta} x')^N} \right) \\ &\quad \times \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x' \hat{\delta} (1-x')} x'^{IN} (1-x')^{N-1} dx'. \end{aligned} \quad (7.30)$$

Theorem 1 shows

$$\begin{aligned} \prod_{i=1}^I \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) &\leq \lim_{\gamma \rightarrow \infty} \prod_{i=1}^I \left(\frac{P \left[\sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} x' \right]}{(\hat{\delta} x')^N} \right) \\ &\leq \prod_{i=1}^I \prod_{n=1}^N \left(\lambda_{i,n} + \mu_{i,n} \right). \end{aligned} \quad (7.31)$$

Since $\lim_{\gamma \rightarrow \infty} e^{-\lambda_0 x' \hat{\delta}(1-x')} = 1$ with $\hat{\delta} = \frac{2^{2R}-1}{\gamma}$, and $\int_0^1 x'^I N(1-x')^{N-1} dx' = \frac{(N-1)!IN!}{(IN+N)!}$, we obtain the following upper and lower bounds with similar derivations

$$L_{AvgBRS} = \frac{\lambda_0^N (IN)!}{(IN+N)!} \prod_{i=1}^I \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) \hat{\delta}^{(I+1)N} \quad (7.32)$$

$$U_{AvgBRS} = \frac{\lambda_0^N (IN)!}{(IN+N)!} \prod_{i=1}^I \prod_{n=1}^N \left(\lambda_{i,n} + \mu_{i,n} \right) \hat{\delta}^{(I+1)N}. \quad (7.33)$$

7.5.4 Per-Subcarrier Best Relay Selection Scheme (NBRS)

For this scheme, the achieved capacity can be written as

$$C_{NBRS} = \frac{1}{2} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \sum_{n=1}^N \max_{i \in [1..I]} \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right), \quad (7.34)$$

and the outage probability is written

$$P_{out}^{NBRS} = P \left[\frac{1}{2} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \sum_{n=1}^N \max_{i \in [1..I]} \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right) < R \right]. \quad (7.35)$$

The outage bounds are derived using the inequalities below

$$\begin{aligned} \max_{\substack{i \in [1..I] \\ n \in [1..N]}} \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} &\leq \sum_{n=1}^N \max_{i \in [1..I]} \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \\ &\leq N \times \max_{\substack{i \in [1..I] \\ n \in [1..N]}} \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}}. \end{aligned} \quad (7.36)$$

The lower bound in (7.36) gives the outage upper bound, and vice versa. By using the limit derived in [22],

$$\lim_{\gamma \rightarrow \infty} \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} = \lambda_{i,n} + \mu_{i,n}, \quad (7.37)$$

and with similar derivations, we obtain

$$L_{NBRS}^{Ref} = \frac{\lambda_0^N (IN)!}{(IN+N)! N^{IN+1}} \prod_{i=1}^I \prod_{n=1}^N \left(\lambda_{i,n} + \mu_{i,n} \right) \hat{\delta}^{(I+1)N} \quad (7.38)$$

$$U_{NBRS}^{Ref} = L_{NBRS}^{Ref} \times N^{IN+1}. \quad (7.39)$$

These bounds will be referred to as *reference bounds* since they are derived from the limit (7.37) from [22], using a similar approach as in [57]. We can also see from (7.19), (7.32) and (7.38) that the APN, AvgBRS and NBRS schemes all achieve full diversity order of $(I+1)N$.

7.5.5 Random Relay Selection Scheme (RRS)

The outage probability for the RRS scheme can be written

$$\begin{aligned} P_{out}^{RRS} &= P \left[\frac{1}{2} \log_2 \left(1 + \sum_{n=1}^N \gamma \alpha_{0,n} + \text{rand}_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \gamma \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \right) < R \right] \\ &= P \left[\sum_{n=1}^N \alpha_{0,n} + \text{rand}_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} \right] \end{aligned} \quad (7.40)$$

with $\hat{\delta}$ defined in (7.28). rand is the function which uniformly picks a random discrete value in $[1..I]$. In the same way as previously, we can write

$$\begin{aligned} P_{out}^{RRS} &= \int_0^1 P \left[\text{rand}_{i \in [1..I]} \sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} x' \right] \\ &\quad \times \hat{\delta}^N \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x' \hat{\delta} (1-x')} (1-x')^{N-1} dx' \end{aligned} \quad (7.41)$$

With the random relay selection, a relay among I is selected in each time frame. Thus, each relay has an equal probability to be selected of $p_i = 1/I$, which is the uniform distribution. As the Probability Mass Function (PMF) $p_m(y)$ of a discrete random variable $y \in [1..I]$ following a uniform distribution is $p_m(y) = \sum_{i=1}^I \frac{1}{I} \times \delta(y-i)$, we can write

$$\begin{aligned} P_{out}^{RRS} &= \int_0^1 \sum_{i=1}^I \frac{1}{I} \times P \left[\sum_{n=1}^N \frac{\gamma \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} < \hat{\delta} x' \right] \\ &\quad \times \hat{\delta}^N \frac{\lambda_0^N}{(N-1)!} e^{-\lambda_0 x' \hat{\delta} (1-x')} (1-x')^{N-1} dx' \end{aligned} \quad (7.42)$$

Finally, using Theorem 1, and with similar derivations, we obtain the bounds

$$L_{RRS} = \frac{\lambda_0^N N}{(2N)!} \sum_{i=1}^I \frac{1}{I} \prod_{n=1}^N \left(\frac{\lambda_{i,n} + \mu_{i,n}}{2} \right) \hat{\delta}^{2N} \quad (7.43)$$

$$U_{RRS} = \frac{\lambda_0^N N}{(2N)!} \sum_{i=1}^I \frac{1}{I} \prod_{n=1}^N \left(\lambda_{i,n} + \mu_{i,n} \right) \hat{\delta}^{2N}. \quad (7.44)$$

From these expressions, we can see that the RRS scheme achieves a diversity order of $2N$ at most, coming from the N subcarriers of the direct path and the N subcarriers of the chosen relay. However, not more than a order N diversity can be guaranteed for the relayed path since the relay is chosen randomly.

7.6 Implementation Issues

The schemes presented above require different amount of CSI feedback. For NBRS and AvgBRS schemes, we assume that the relay selection is performed at the destination, based on its received SNR in each subcarrier. To this end, all the relays need to send pilots over all subcarriers using some dedicated OFDM control symbols in the DL frame. Then, the destination provides feedback of the required information to the source via an uplink channel. For NBRS scheme, the destination feeds back the ID of the selected relay for each subcarrier, which makes $N \times n_{ID}$ bits where n_{ID} is the number of bits used for describing the identity of each relay. In AvgBRS scheme, the destination feeds back to the source only one relay ID, e. g., n_{ID} bits. Finally, the APN scheme, AP-RS scheme as well as RRS scheme do not require any feedback, all or one of the relays send blindly their received signal to the destination.

We expect the performance of NBRS to be the best among all the schemes since the adaptation is made per subcarrier. But from a practical point of view, we can think that this scheme will be hardly realizable since it requires a large amount of feedback, especially when N is high. Also, when we consider a multi-user setting, it will become too prohibitive if all the users send this feedback. Another problem is that all the transmissions from the different relays in each subcarrier have to be synchronized over time and frequency, otherwise the signal will not be successfully decided at the destination. However, this is very difficult to achieve when different relays are transmitting in different subcarriers. Considering these implementation matters, AvgBRS seems more feasible since it only requires to feedback one relay ID per user and only one relay transmits, which avoids such synchronization problems.

7.7 Numerical Results

7.7.1 Simulation results and comparison with the theoretical analysis

First, we consider a SC system, to simply illustrate the results of Theorem 1 from section 7.4. Using the limit in (7.7) derived in [22], [57] showed that in the high SNR regime, the outage probability of the AP-AF scheme P_{out}^{APAF} can be bounded by L_{APAF} and U_{APAF} defined as

$$\begin{aligned} L_{APAF} &= \frac{\lambda_0 \prod_{i=1}^I (\lambda_i + \mu_i)}{(I+1)I^I} \left(\frac{2^{(I+1)R} - 1}{\gamma} \right) \\ U_{APAF} &= L_{APAF} \times I^I \end{aligned} \quad (7.45)$$

Using Theorem 1, and with similar derivations, we obtain the following bounds

$$\begin{aligned} L'_{APAF} &= \frac{\lambda_0}{(I+1)} \prod_{i=1}^I \left(\frac{\lambda_i + \mu_i}{2} \right) \left(\frac{2^{(I+1)R} - 1}{\gamma} \right) \\ U'_{APAF} &= U_{APAF}. \end{aligned} \quad (7.46)$$

Simulations are made for $N = 1$ subcarrier, $I = 3$ relays and R fixed to 1 [b/s/Hz]. Fig. 7.7 shows that our derived lower bound L'_{APAF} is much tighter than L_{APAF} . L'_{APAF} can be interpreted as a tight approximation of the AP-AF outage, for SNR values higher than 12dB. Similarly, we can expect that our derivations using Theorem 1 provide a tighter lower bound in the MC setting than the reference bounds.

For assessing our analysis made in the MC case, we also define reference bounds using the approach in [57] for the AvgBRS and APN schemes. From the outage expression of AvgBRS in (7.27), we can see that the left and right bounds for the NBRS scheme in (7.36) are also valid for the AvgBRS scheme. Thus, AvgBRS and NBRS share the same reference bounds (7.38), denoted L_{BRS}^{Ref} and U_{BRS}^{Ref} for simplicity, so that

$$\begin{aligned} L_{BRS}^{Ref} &= L_{AvgBRS}^{Ref} = L_{NBRS}^{Ref} \\ U_{BRS}^{Ref} &= U_{AvgBRS}^{Ref} = U_{NBRS}^{Ref}. \end{aligned} \quad (7.47)$$

In the same way, using the inequalities

$$\begin{aligned} \max_{\substack{i \in [1..I] \\ n \in [1..N]}} \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} &\leq \sum_{i=1}^I \sum_{n=1}^N \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \\ &\leq IN \times \max_{\substack{i \in [1..I] \\ n \in [1..N]}} \frac{\gamma^2 \alpha_{i,n} \beta_{i,n}}{1 + \gamma \alpha_{i,n} + \gamma \beta_{i,n}} \end{aligned} \quad (7.48)$$

and (7.37), the following reference bounds can be derived for the APN scheme

$$L_{APN}^{Ref} = \frac{\lambda_0^N (IN)!}{(IN+N)!(IN)^{IN}} \prod_{i=1}^I \prod_{n=1}^N (\lambda_{i,n} + \mu_{i,n}) \tilde{\delta}^{(I+1)N} \quad (7.49)$$

$$U_{APN}^{Ref} = L_{APN}^{Ref} \times (IN)^{IN} \quad (7.50)$$

Comparing the analytical expressions in (7.49) and (7.20), we can see that the reference upper bound U_{APN}^{Ref} and the upper bound U_{APN} are equal. For AvgBRS,

we also have $U_{BRS}^{Ref} = U_{AvgBRS}$ (see (7.32) and (7.49)). Therefore, only U_{APN}^{Ref} and U_{BRS}^{Ref} are represented in the figures.

In the MC case, the simulations were made assuming $I=2$, $N=2$, and the target rate R was fixed to 2 bits per frame (equivalent to 1 bit per frame, per subcarrier). Rayleigh fading channels with parameters $h_{s,d,n}, h_{s,i,n}, h_{i,d,n} \sim CN(0, 1)$, $\forall(i, n)$ are considered, and thus the parameters of the exponential random variables $\alpha_{0,n}$, $\alpha_{i,n}$ and $\beta_{i,n}$ verify $\lambda_0 = \lambda_{i,n} = \mu_{i,n} = 1$, $\forall(i, n)$. Fig. 7.8 shows the performance of the four schemes APN, AvgBRS, NBRS and RRS. We can observe that the all-participate scheme APN performs very poorly compared to the best relay selection schemes, AvgBRS and NBRS, as it was the case for the SC system studied in [57]. This is due to the equal time partition between all the access points which degrades the throughput and thus the outage probability. It is also shown that the outage of AvgBRS, is extremely close to the performance of NBRS. AvgBRS provides a near-optimal performance while drastically decreasing the required amount of CSI, since it only needs the feedback of I CSI values, e. g., the average SNR per relay, whereas NBRS needs the feedback of $N \times I$ CSI values, e. g., the SNR for each subcarrier and relay. This can be explained by the fact that, since NBRS can only choose between 2 relays per subcarrier, the choice will not differ so much compared to AvgBRS, as the best relay in average is likely to be chosen in some of the subcarriers for NBRS. We can expect that, as the number of relays grows, the gap between the two schemes may increase, as we will see later. However, compared to the RRS scheme, AvgBRS achieves a much better performance with a low amount of CSI. Thus, AvgBRS is very well suited for practical use. Moreover, we can see that the lower bound derived for the AvgBRS scheme, L_{AvgBRS} , improves its reference lower bound L_{BRS}^{Ref} and that the lower bound of RRS L_{RRS} approaches well the outage probability of RRS.

Fig. 7.9 shows the reference and derived bounds for the APN and AP-RS schemes, for different values of the average SNR, common to all links. For the AP-RS scheme, the bounds L_{APRS} and U_{APRS} are not plotted for low SNR values since they have negative values (see (7.24) and (7.25)). The theoretical bounds match well the simulated performance of AP-RS. We can observe that the APN scheme performs much better than the AP-RS scheme, except for SNR values lower than 15dB, where the AP-RS scheme performs slightly better. The simulated curves and the derived bounds (7.19), (7.20) show that coding the data across all the subcarriers provides full diversity, which is not the case with rate splitting. This can be understood by the fact that if one subcarrier supports a rate which is lower than R/N , then the whole scheme is in outage, e. g., there is no diversity gain from the subcarriers, which explains the poor outage performance. Concerning the APN scheme, our derived lower bound L_{APN} is much tighter than the reference bound L_{APN}^{Ref} and provides a very good approximation of the APN

outage. However, we can observe that L_{APN} becomes a lower bound for the APN outage only for SNR higher than 22dB. This is due to the power $I + 1$ in the expression of $\tilde{\delta}$ which makes it tend to zero (e. g., γ tends to infinity) at slower speed than $\hat{\delta}$, so that higher SNRs are required for the bounds of APN to be valid.

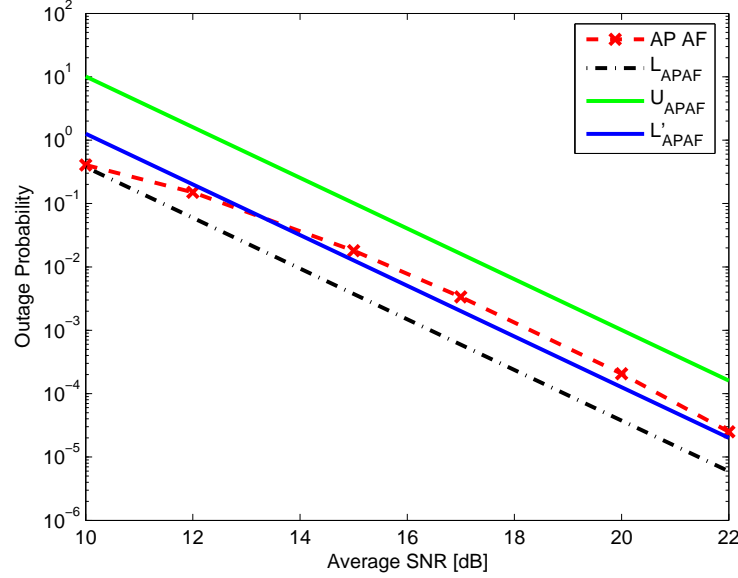


Figure 7.7: Outage Probability and bounds, AP-AF scheme ($N=1$, $I=3$, $R=1$)

7.7.2 Simulation results assuming a practical OFDM channel model

In the previous part, we have shown that the AvgBRS scheme approached the outage performance of the NBRS scheme, while considerably decreasing the required amount of CSI, when we assume the subcarriers as independent Rayleigh fading channels. Now, we verify this result for a practical OFDM channel, generated according to the multipath Rayleigh fading channel model with exponential power delay profile used in [50]. We evaluated the Cumulative Density Function (CDF) of the capacity for the APN, AvgBRS, NBRS and RRS schemes, with $N = 48$ subcarriers and $I = 2$ relays, for a fixed average SNR level of $15dB$ in all links. Fig. 7.10 shows that in this practical case too, AvgBRS approaches the performance of NBRS, and both schemes achieve a considerably higher performance compared to the APN scheme.

By increasing the number of relays to $I = 12$, we can see from Fig. 7.11 that the gap between NBRS and AvgBRS is increased. As the number of relays increases,

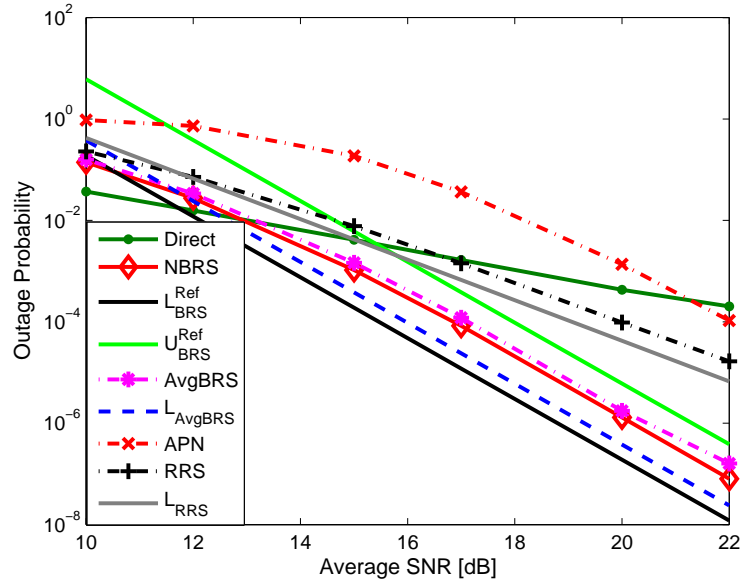


Figure 7.8: Outage Probability and bounds for the proposed schemes

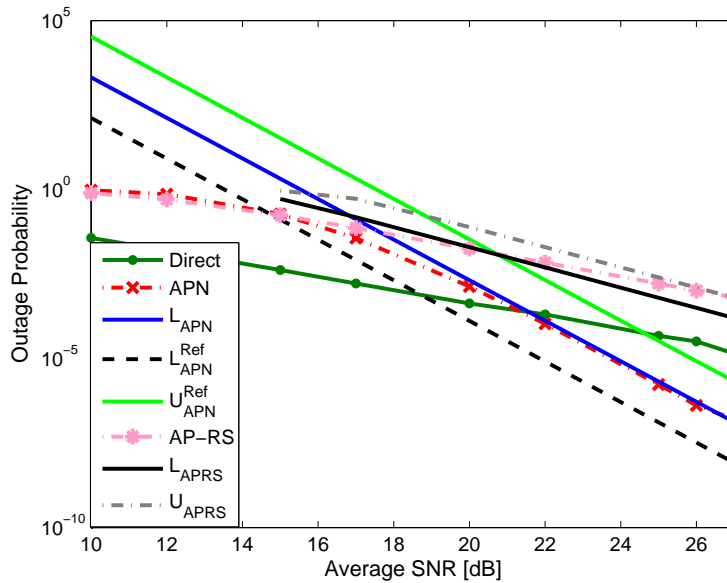


Figure 7.9: Outage Probability and bounds for APN and AP-RS Schemes

there is a higher probability that a different relay is chosen in each subcarrier by the NBRs scheme, compared to the relay allocated by the AvgBRS scheme in all subcarriers. But we can observe that it is the capacity of NBRs which is improved, while AvgBRS achieves a constant performance in both cases. Moreover, the performance achieved by the APN scheme is much degraded due to the orthogonal partition of the time frame among the I relays.

In both cases, when $I = 2$ and $I = 12$, for a large number of subcarriers ($N = 48$), the performance of RRS and AvgBRS are comparable. This is because as the number of subcarriers increases, the average SNR achieved in each relay becomes comparable, so the gap in the SNR averaged over the subcarriers between the best relay and the worst one is minimized. However, we know from the theoretical analysis in section 7.5 that AvgBRS achieves a full diversity order of $(N + 1)I$ whereas RRS achieves only a $2N$ diversity order. Thus, AvgBRS seems the best scheme for practical use when considering OFDM, as it achieves the full diversity order and approaches the capacity performance of NBRs with a minimal amount of CSI feedback. Moreover, in this simulation, the same average SNR was assumed for all links, but we can expect the gain of the AvgBRS scheme over the RRS scheme to increase when each relay–destination link achieves a different average SNR level.

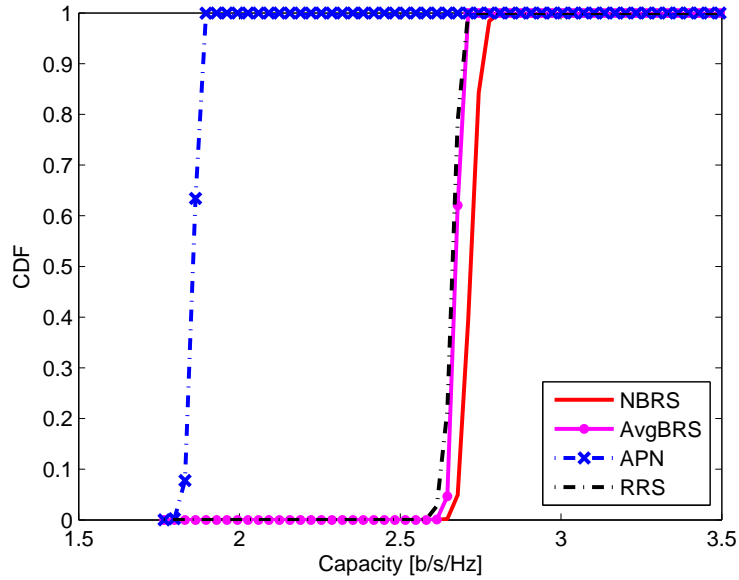


Figure 7.10: CDF of the capacity [b/s] of the proposed schemes, OFDM channel $N = 48$, $I = 2$

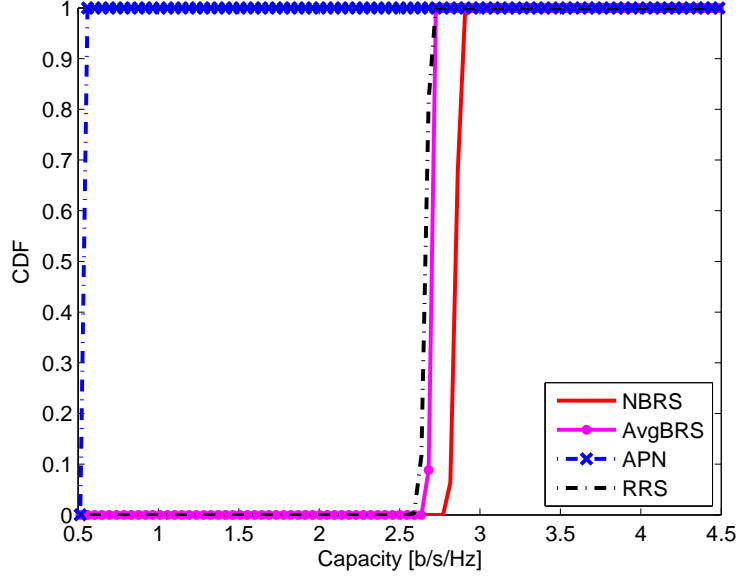


Figure 7.11: CDF of the capacity [b/s] of the proposed schemes, OFDM channel $N = 48$, $I = 12$

7.8 Summary

In this chapter, we have investigated cooperative diversity schemes using the AF protocol for the MC wireless communication system. Assuming multiple relays, several allocation schemes have been proposed. For each scheme, theoretical bounds of the outage probability have been derived via analysis using our generalization of the limit to the multiple link case. The simulations have shown the improvement of our bounds compared to reference bounds. In particular, the lower bound for the All-Participate All Subcarriers (APN) scheme provides an excellent approximation. Moreover, spreading the data accross all subcarriers as in the APN scheme gave a better outage than splitting it into independent streams as in the All-Participate Rate Splitting (AP-RS) scheme, especially in the high SNR region. The results showed that the Average Best Relay Selection (AvgBRS) scheme had a near-optimal performance as it closely approached the outage of NBRs scheme, while achieving full diversity order as opposed to the Random Relay Selection (RRS) scheme. This was also observed when considering a practical OFDM channel model. This is very interesting from the practical point of view, since AvgBRS requires N times less CSI feedback compared to NBRs, for a similar performance.

*CHAPTER 7. AMPLIFY-AND-FORWARD COOPERATIVE DIVERSITY
SCHEMES FOR MULTI-CARRIER SYSTEMS*

Chapter 8

Conclusions and Future Work

This thesis has focused on the radio resource allocation and scheduling issues in the next generation wireless communication system assuming Multi-Carrier (MC) transmission technology. The study has targeted the design of such algorithms for two different system architectures:

1. Cellular MC system
2. Relay-based MC system.

The main conclusions from this work can be summarized around the following main axes.

8.1 Resource allocation and Scheduling in Cellular MC system

OFDMA systems offer the possibility to achieve high system performance by exploiting the multi-user diversity effect over time and frequency. However, providing user fairness and QoS satisfaction are also essential aspects of the scheduling problem. Designing such efficient algorithms with low complexity for OFDMA systems is more challenging due to the additional degree of freedom in frequency. We have focused on the PFS algorithm which gave an excellent trade-off between throughput and fairness, in Single-Carrier (SC) systems. However, the optimal algorithm for achieving PF in MC systems requires a prohibitively high complexity, which makes it very difficult to implement in real systems. On the other hand, suboptimal PFS algorithms for MC systems proposed in the literature achieve very low levels of throughput and do not fully exploit the MC property. We have designed a number of suboptimal algorithms based on PFS, which provide a very

good throughput/fairness trade-off, improving the performance of existing algorithms, considering the following 2 system assumptions.

8.1.1 Per-Slot PFS in a MC System

We have first focused on the case of slot-by-slot scheduling, which is the most widespread assumption but hardly the most practical one. The proposed algorithms, *Best-PFS* and *Best-Rate*, improved both throughput and PF, compared to the *Conventional MC-PFS* algorithm, where SC PFS algorithm is applied per subchannel. The improvement is obtained by a joint optimization of PFS and rate, by prioritizing the users for which both PFS and rate metrics are the best. Concerning the *Best-Rate* algorithm, a large increase in PF can be observed compared to the *Max CSI* which allocates the user with the best rate in each algorithm. This is due to the fact that, with discrete AMC levels, many users have actually the same achievable rate, and PF is increased by choosing the user with the highest PFS metric among all the best-rate users at the same level. Moreover, our algorithms provide a deterministic subchannel ordering, as opposed to most of the existing algorithms characterized by arbitrary ordering. We developed a method which provides the bounds of the optimal PFS, which would be otherwise too complex to compute for the considered number of users and subchannels. The results showed that *Best-PFS* achieved a near-optimal PF, and *Best-Rate*, a near-optimal throughput.

8.1.2 Multi-Slot PFS in a MC System

In the practical case, such as in the standardized IEEE 802.16 system, a frame is composed of multiple time slots and scheduling should be performed for all these slots simultaneously rather than slot-by-slot. Existing scheduling algorithms allocate a unique user per subchannel, e. g., they do not allow that different users are allocated different time slots in one subchannel, during a frame. Exploiting the degree of freedom in time jointly with the subchannel allocation may yield significant improvement in terms of fairness, since a higher number of users may be multiplexed into each frame with the increased granularity of the allocation. We have first investigated the conditions for achieving optimal PF when the frame is composed of multiple slots, and derived its bounds. 2 suboptimal algorithms were designed, *PF-Mux* and *Rate-Mux*. While significant PF gain can be achieved with user multiplexing, allocating different users on different slots increases the complexity for describing this scheduling decision, thereby increasing the amount of signalling for encoding the user mapping. This drawback was mitigated in our proposed algorithms by introducing some predefined patterns that user multiplexing should follow, hence considerably decreasing the amount of required

signalling. Thus, the algorithms were designed in such a way that PF and rate maximization is provided, while minimizing the signalling overhead. A very accurate estimation of the optimal PF is achieved by our bounding approach. With low complexity, our proposed algorithms provided an excellent trade-off in terms of throughput, PF and starvation time, defined as the maximum time that a user does not receive service. Especially, comparison with the *Mux MC-PFS* algorithm (simple application of the *Conventional MC-PFS* algorithm per slot) which resulted in an extremely poor throughput due to the increased signalling, shows the real benefit of our algorithms.

There are several interesting continuations of this work:

- Performance of user multiplexing algorithms with imperfect CSI: imperfectness of the CSI increases with user mobility, as the higher the user speed, the higher the channel variations and the more outdated the CSI. Multiplexing different users in each subchannel may improve transmission efficiency since one subchannel would not be completely lost if the CSI of one user becomes totally erroneous. However, when the CSI of all users become unreliable, for example when all users experience very high mobility like in trains, there would not be any advantage of our proposed schemes compared to systematic frequency/time-hopping schemes. It is worth investigating on the upper limit of the user speed for which our algorithms still provide some gain.
- In the case of imperfect CSI and heterogeneous user mobilities, the user order in the allocated time slots may impact on the overall performance. For example, users with higher mobility may be allocated the first slots where the CSI is still reliable, while users with lower mobility can be allocated the last slots. The impact of such allocations should be evaluated. But this requires to encode the allocated user order over the time slots, thereby increasing the amount of signalling. Some adaptation of the algorithms would be needed.
- The integration of the proposed algorithms in a Multiple Input Multiple Output (MIMO) system may further improve the system performance.

8.1.3 Adaptive CSI provision in a MC System

Providing user CSI to the BS scheduler is a crucial problem in a MC system. While frequent and substantial amount of CSI is required to ensure a good scheduling, UL channel resource is wasted to carry this information. It is particularly problematic in a MC system which comprises a large number of subcarriers and user channel quality can highly vary over the subcarriers. We have first observed

that depending on the scheduler and the number of users in the cell, the amount of useful CSI varies, since the only useful CSI are the ones belonging to the users who will be served. That is, the CSI required by *Max CSI* and by a *PFS* differs, as the users who are scheduled are not the same. The number of users supporting the best rate increases due to the multi-user diversity effect, so only their CSI needs to be reported for *Max CSI*. Thus, we have developed a method for CSI provision which can adapt to these variable circumstances and optimize its message length according to the degree of CSI accuracy required by the scheduler. We also proposed a *Differential Encoding* scheme which is very efficient compared to existing methods. Simulation results have shown that a significant feedback reduction was achieved with our proposed scheme, while keeping a good scheduling performance, especially for *Max CSI*. While feedback reduction is more limited for *PFS*, since more CSI needs to be reported in order to keep PF, substantial reduction was obtained with our scheme. This scheme is a generic one which can be used for various types of schedulers, not limited to *Max CSI* and *PFS*.

Several directions for improvement can be pointed out:

- While the proposed scheme can adapt to the BS decisions on the amount of CSI to be reported, how this amount is decided remains a difficult issue as it depends on many factors. Thorough investigation may be worthwhile.
- Provision of CSI becomes even more problematic when relays come into play, since the number of links is multiplied. Design of efficient encoding schemes is a big issue.
- Integration of the proposed encoding in a MIMO system.

8.2 Resource allocation and Scheduling in Relay-based MC system

The impact of relaying on resource allocation and scheduling was considered in the second part of the thesis. Problems concerning path selection, frame design, subchannel/time allocation and CSI provision have been addressed. Feasible frames and allocation options were derived taking into account the practical limitations imposed by the system. Algorithms for resource allocation have been proposed, for single RS and multiple RS cases. In order to decrease the algorithm complexity and the amount of required CSI, allocation of relayed users is decoupled from the direct users. Each RS makes its own initial allocation, and forwards this decision to the BS, which optimizes the overall allocation. A tremendous decrease of CSI feedback is achieved since the CSI of relayed users is only needed

at the corresponding RS and not at the BS. Moreover, the designed algorithms are based on a modified PFS criterion for ensuring a minimal per-user throughput. For performance assessment, we have designed 2 optimal algorithms, one for throughput and the other for outage, where resource assignment is no longer constrained by the above mentioned practical limitations.

In the single RS case, two algorithms have been designed: the *FTD* algorithm operating with fixed time division and the *ATD* algorithm, with time division optimization. These algorithms were evaluated assuming two different user distributions; the uniform distribution, and distribution towards the cell edge. While the performance enhancement of the proposed algorithms was limited in the case of uniform user distribution, both algorithms outperformed the *Conventional MC-PFS* algorithm performed without relays when users were distributed towards the cell edge, which is the typical scenario targeted for making use of relays. With the time adaptation, *ATD* improved the performance of *FTD* at the cost of reasonably higher complexity. A new operation was introduced in our algorithms, where the RS requests the BS to forward new packets for the relayed users who are likely to be scheduled over the upcoming frames. Compared to the case where the BS forwards all the packets for relayed users at any time, our scheme provided tremendous gain in throughput/outage, since it realized optimal resource distribution in the BS-subframe between the BS-RS links and the direct users.

In the multiple relay case, various allocation options have been considered with static or adaptive time divisions. This time, the performance enhancement brought by the proposed algorithms compared to the *Conventional MC-PFS* were also notable in the case of uniform user distribution, as well as when users were distributed towards the cell edge. The comparison of the static time division and adaptive division schemes showed that a significant gain in throughput/outage performance could be achieved. The adaptation is based on the concept of RS activation, where the frame structure is adapted according to the active RS, thereby increasing spectrum utilization. The *MRP* algorithm which is the static allocation, and the *MRPA* algorithm, a basic RS activation step which allocates the frame only to the relays where users are attached, can significantly improve the outage performance with low complexity. By further selecting the most efficient relays among the activated ones, *MRAA* increases the achieved throughput while keeping the outage at low level. Adapted to various number of relays, the proposed algorithms and especially *MRAA* always achieved a good trade-off between outage and throughput compared to reference algorithms, while minimizing the computational complexity and required amount of CSI.

There are several possible further developments of these algorithms:

- The proposed algorithms may be adapted for multi-hop relaying systems, where a relayed link may contain more than 2 hops.

- Combining the proposed algorithms with space–time coding techniques with MIMO may further reduce the system outage and increase throughput.

8.3 Cooperative Diversity Schemes for a relay–aided MC System

In the last part of the thesis, cooperative diversity schemes were proposed using the AF protocol for the MC wireless communication system. 4 types of schemes were proposed and analyzed. Since the obtention of the exact outage probability is mathematically intractable, upper and lower bounds were derived by means of approximation, a method specifically developed for the MC/multiple relay case. The lower bound obtained with our method is much tighter than the one derived from a simple extension of the SC lower bound to the MC case, which shows the efficiency of our approach. The comparison of the proposed schemes using the theoretical bounds showed that selecting the best relay per subcarrier as in the *NBRS* scheme gives the best outage performance but requires tremendous CSI feedback, while splitting the frame between all the relays as in the *APN* scheme is very inefficient. The *AvgBRS* scheme, where the best relay in average over all subcarriers is allocated, makes the best trade–off as it achieves a near–optimal outage while minimizing the required feedback. It is the most suited scheme for practical implementation. These results were confirmed when considering a practical OFDM system.

These results open the field for interesting future issues such as:

- Design of such allocation schemes when there are multiple users and theoretical analysis of their achievable performance.
- When there are multiple users, how to deal with the explosion of the amount of CSI feedback.
- Consideration of the MAC layer issues, such as ARQ schemes.
- Extension to MIMO.

Appendix A

List of Publications produced during the Ph.D. Studies

The dissemination of the results obtained within the Ph.D. work has been targeted to international conferences and journals. The following list classifies the publications in relation to the respective chapters of the thesis.

Publications related to Chapter 3:

- **Journal:**

1. M. Kaneko, P. Popovski and J. Dahl, "Proportional Fairness in Multi-Carrier System: Upper Bound and Approximation Algorithms", *IEEE Communication Letters*, Vol. 10, No. 6, pp. 462-464, June 2006

Publications related to Chapter 4:

- **Journal:**

1. M. Kaneko, P. Popovski and J. Dahl, "Proportional Fairness in Multi-Carrier System with Multi-Slot Frames: Upper Bound and User Multiplexing Algorithms", *accepted for publication in IEEE Transactions on Wireless Communications*, Jan. 2007

- **Conference:**

1. M. Kaneko and P. Popovski, "Heuristic Subcarrier Allocation Algorithms with Multi-Slot Frames in Multi-User OFDM Systems", *Proc. IEEE International Conference on Communications (ICC)*, Istanbul, Turkey, June 2006

APPENDIX A. LIST OF PUBLICATIONS PRODUCED DURING THE PH.D.
STUDIES

Publications related to Chapter 5:

• **Conference:**

1. M. Kaneko, P. Popovski and H. Yomo, "Adaptive Provision of CSI Feedback in OFDMA Systems", *Proc. IEEE 17th Personal, Indoor and Mobile Radio Communications (PIMRC)*, Helsinki, Finland, Sept. 2006

Publications related to Chapter 6:

• **Journal:**

1. M. Kaneko, P. Popovski and K. Hayashi, "Throughput-Guaranteed Resource Allocation Algorithms for Relay-aided Cellular OFDMA System", *Submitted to IEEE Transactions on Vehicular Technology*, June 2007

• **Conference:**

1. M. Kaneko and P. Popovski, "Adaptive Resource Allocation in Cellular OFDMA System with Multiple Relay Stations", *Proc. IEEE 65th Vehicular Technology Conference (VTC-Spring)*, Dublin, Ireland, April 2007
2. M. Kaneko and P. Popovski, "Radio Resource Allocation Algorithm for Relay-aided Cellular OFDMA System", *Proc. IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, June 2007

Publications related to Chapter 7:

• **Journal:**

1. M. Kaneko, K. Hayashi, P. Popovski, K. Ikeda, H. Sakai and R. Prasad, "Amplify-and-Forward Cooperative Diversity Schemes for Multi-Carrier Systems", *Conditionally accepted in IEEE Transactions on Wireless Communications, with major revision*, June 2007

Other Publications:

• **Patents:**

1. P. Popovski, M. Kaneko, J.M. Ro, E.T. Lim, Y.K. Cho and D.S. Park, "Method and System for Transmitting/Receiving Data in Communication System", KR2006-0004984, filed in Korean Patent Office, and US678-2930, filed in US Patent Office, Jan. 2006

APPENDIX A. LIST OF PUBLICATIONS PRODUCED DURING THE PH.D. STUDIES

2. M. Kaneko, P. Popovski, Y. Chang, E.T. Lim, P. Joo, C. Oh., C. Shan and Y.H. Park, "Apparatus and Method for Path Selection in Broad-band Wireless Communication System", KR2007-0017821, filed in Korean Patent Office, 22nd February 2007
 3. M. Kaneko, P. Popovski, Y. Chang, E.T. Lim, P. Joo, C. Oh., C. Shan and Y.H. Park, "Apparatus and Method for Resource Allocation considering Buffering in Relay Wireless Communication System", KR2007-0017699, filed in Korean Patent Office, 22nd February 2007
- **Project Deliverables:**
1. M. Kaneko, N. Marchetti, D. Figueiredo, P. Popovski and F. Fitzek, "Multi-User Diversity in Scheduling", JADE Project D.3.2.2, Jan. 2005
 2. M. Kaneko and P. Popovski, "Radio Resource Allocation in DL OFDMA: July-December 2005", JADE Project, Dec. 2005
 3. M. Kaneko and P. Popovski, "Resource Allocation and Scheduling for Relay-aided Cellular OFDMA System", JADE Project, June 2006
 4. M. Kaneko and P. Popovski, "Radio Resource Allocation and Scheduling for Relay-aided Cellular OFDMA System: July-November 2006", JADE Project, Nov. 2006

*APPENDIX A. LIST OF PUBLICATIONS PRODUCED DURING THE PH.D.
STUDIES*

Appendix B

Proof of Theorem 1

Here we give the proof for $N=2$. The bounds for $N > 2$ can be easily obtained via mathematical induction. We consider independent exponential random variables x, y, z and t with parameters $\lambda_x, \lambda_y, \lambda_z, \lambda_t$, respectively. We observe

$$\begin{aligned} r_\delta &= \frac{1}{\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy}} + \frac{1}{\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt}} \\ &\leq \frac{1}{\frac{1}{x} + \frac{1}{y}} + \frac{1}{\frac{1}{z} + \frac{1}{t}} \leq 2 \max \left(\frac{1}{\frac{1}{x} + \frac{1}{y}}, \frac{1}{\frac{1}{z} + \frac{1}{t}} \right), \end{aligned} \quad (\text{B.1})$$

Therefore, we have

$$\begin{aligned} P[r_\delta < h(\delta)] &\geq P \left[2 \max \left(\frac{1}{\frac{1}{x} + \frac{1}{y}}, \frac{1}{\frac{1}{z} + \frac{1}{t}} \right) < h(\delta) \right] \\ &= P \left[\min \left(\frac{1}{x} + \frac{1}{y}, \frac{1}{z} + \frac{1}{t} \right) > \frac{2}{h(\delta)} \right] \\ &\geq P \left[\min \left(\max \left(\frac{1}{x}, \frac{1}{y} \right), \max \left(\frac{1}{z}, \frac{1}{t} \right) \right) > \frac{2}{h(\delta)} \right]. \end{aligned} \quad (\text{B.2})$$

Since $\frac{1}{x}, \frac{1}{y}, \frac{1}{z}$ and $\frac{1}{t}$ are independent random variables,

$$\begin{aligned}
 P[r_\delta < h(\delta)] &\geq P\left[\max\left(\frac{1}{x}, \frac{1}{y}\right) > \frac{2}{h(\delta)}\right] P\left[\max\left(\frac{1}{z}, \frac{1}{t}\right) > \frac{2}{h(\delta)}\right] \\
 &= \left(1 - P\left[\max\left(\frac{1}{x}, \frac{1}{y}\right) \leq \frac{2}{h(\delta)}\right]\right) \left(1 - P\left[\max\left(\frac{1}{z}, \frac{1}{t}\right) \leq \frac{2}{h(\delta)}\right]\right) \\
 &= \left(1 - P\left[\min(x, y) \geq \frac{h(\delta)}{2}\right]\right) \left(1 - P\left[\min(z, t) \geq \frac{h(\delta)}{2}\right]\right) \\
 &= \left(1 - P\left[x \geq \frac{h(\delta)}{2}\right] P\left[y \geq \frac{h(\delta)}{2}\right]\right) \times \left(1 - P\left[z \geq \frac{h(\delta)}{2}\right] P\left[t \geq \frac{h(\delta)}{2}\right]\right) \\
 &= \left(1 - \exp\left(-\frac{\lambda_x + \lambda_y}{2} h(\delta)\right)\right) \left(1 - \exp\left(-\frac{\lambda_z + \lambda_t}{2} h(\delta)\right)\right).
 \end{aligned} \tag{B.3}$$

Moreover, from Fact 1 of [22], we know

$$\lim_{\delta \rightarrow 0} \frac{1 - \exp\left(-\frac{\lambda_i + \lambda_j}{2} \times h(\delta)\right)}{h(\delta)} = \frac{\lambda_i + \lambda_j}{2}, \tag{B.4}$$

which finally leads to the lower bound

$$\lim_{\delta \rightarrow 0} \frac{1}{h^2(\delta)} P[r_\delta < h(\delta)] \geq \left(\frac{\lambda_x + \lambda_y}{2}\right) \left(\frac{\lambda_z + \lambda_t}{2}\right). \tag{B.5}$$

For the upper bound, we observe

$$\begin{aligned}
 r_\delta &= \frac{1}{\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy}} + \frac{1}{\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt}} \\
 &\geq \max\left(\frac{1}{\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy}}, \frac{1}{\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt}}\right),
 \end{aligned} \tag{B.6}$$

which implies

$$\begin{aligned}
 P[r_\delta < h(\delta)] &\leq P\left[\max\left(\frac{1}{\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy}}, \frac{1}{\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt}}\right) < h(\delta)\right] \\
 &= P\left[\min\left(\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy}, \frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt}\right) > \frac{1}{h(\delta)}\right] \\
 &= P\left[\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy} > \frac{1}{h(\delta)}\right] P\left[\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt} > \frac{1}{h(\delta)}\right].
 \end{aligned} \tag{B.7}$$

It was shown in [22] that

$$\begin{aligned} \frac{1}{h(\delta)} P\left[\frac{1}{x} + \frac{1}{y} + \frac{\delta}{xy} > \frac{1}{h(\delta)}\right] &\leq \lambda_x + \lambda_y \\ \frac{1}{h(\delta)} P\left[\frac{1}{z} + \frac{1}{t} + \frac{\delta}{zt} > \frac{1}{h(\delta)}\right] &\leq \lambda_z + \lambda_t. \end{aligned} \tag{B.8}$$

Thus

$$\frac{1}{h^2(\delta)} P\left[r_\delta < h(\delta)\right] \leq (\lambda_x + \lambda_y)(\lambda_z + \lambda_t). \tag{B.9}$$

Bibliography

- [1] R. Prasad and M. Ruggieri, *Technology Trends in Wireless Communications*. London, UK: Artech House Publishers, 2003.
- [2] S. Hara and R. Prasad, *Multicarrier Techniques for 4G Mobile Communications*. London, UK: Artech House Publishers, 2003.
- [3] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. London, UK: Artech House Publishers, 2000.
- [4] “Feasibility Study for OFDM for UTRAN enhancement (Release 6),” 3GPP TR 25.892 v0.5.2., December 2003.
- [5] “IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems,” IEEE Std. 802.16, Rev. 2004, 2004.
- [6] “IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands,” IEEE Std. 802.16e, Rev. 2005, 2005.
- [7] R. Pabst et al., “Relay-Based Deployment Concepts for Wireless and Mobile Broadband Radio,” *IEEE Wireless Comm. Mag.*, pp. 80–89, September 2004.
- [8] “Radio Resource Management and Protocol Engineering for IEEE 802.16,” *IEEE Wireless Commun. Magazine*, Vol. 14, Feb. 2007.
- [9] P. Viswanath and D.N.C Tse and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Trans. Info. Theo.*, vol. 48, no. 6, pp. 1277–1294, June 2002.
- [10] R. Knopp and P. Humblet, “Information capacity and power control in single cell multiuser communications,” in *Proc. IEEE ICC*, vol. 1, Seattle, WA, June 1995, pp. 331–335.

BIBLIOGRAPHY

- [11] J. Jang and K.B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 2, pp. 171–178, February 2003.
- [12] D.N.C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Symp. Information Theory*, Ulm, Germany, June 1997.
- [13] C.Y. Huang et al., "Radio Resource Management of Heterogeneous Services in Mobile WiMAX Systems," *IEEE Wirel. Commun.*, vol. 14, pp. 20–26, February 2007.
- [14] L. Bardia et al., "On the Impact of Physical Layer Awareness on Scheduling and Resource Allocation in Broadband Multicellular IEEE 802.16 Systems," *IEEE Wirel. Commun.*, vol. 14, pp. 36–43, February 2007.
- [15] C. Y. Wong, R. S. Cheng, K.B. Letaief and R.D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, October 1999.
- [16] Qualcomm, "1× EV: 1× Evolution IS-856 TIA/EIA Standard Airlink Overview," *Revision 7.2*, November 2001.
- [17] S. Yoon, Y. Cho, C.B. Chae and H. Lee, "System Level Performance of OFDMA forward link with proportional fair scheduling," in *Proc. IEEE PIMRC*, vol. 2, Barcelona, Spain, September 2004, pp. 1384–1388.
- [18] G. Song and Y. Li, "Cross-Layer Optimization for OFDM Wireless Networks–Part II: Algorithm Development," *IEEE Trans. Wireless Comm.*, vol. 4, no. 2, pp. 625–634, March 2005.
- [19] H. Kim and Y. Han, "A proportional fair scheduling for multicarrier transmission systems," *IEEE Commun. Letters*, vol. 9, no. 3, pp. 210–212, March 2005.
- [20] "Harmonized Contribution on 802.16j (Mobile Multihop Relay) Usage Models," IEEE 802.16j-06/015, September 2006.
- [21] "Recommendations for the Scope and Purpose of the Mobile Multihop Relay Task Group," IEEE 802.16mmr-05/032, November 2005.
- [22] J.N. Laneman, D.N.C. Tse and G.W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Info. Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.
- [23] J. Proakis, *Digital Communications*. McGraw-Hill, 1995.

BIBLIOGRAPHY

- [24] S. T. Chung and A. J. Goldsmith, "Degrees of Freedom in Adaptive Modulation: A Unified View," *IEEE Trans. Comm.*, vol. 49, no. 9, pp. 1561–1571, September 2001.
- [25] S. H. Ali et al., "Dynamic Resource Allocation in OFDMA Wireless Metropolitan Area Networks," *IEEE Wirel. Commun.*, vol. 14, pp. 6–13, February 2007.
- [26] G. Song and Y. Li, "Cross-Layer Optimization for OFDM Wireless Networks—Part I: Theoretical Framework," *IEEE Trans. Wireless Comm.*, vol. 4, no. 2, pp. 614–624, March 2005.
- [27] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *IEEE GLOBECOM*, San Francisco, CA, December 2000.
- [28] C. Y. Wong, C.Y. Tsui, R.S. Cheng and K.B. Letaief, "A real-time subcarrier allocation scheme for multi-access downlink OFDM transmission," in *IEEE VTC*, Ulm, Germany, Fall 1999.
- [29] Y. Zhang and K.B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1566–1575, September 2004.
- [30] W. Anchun, X. Liang, Z. Shidong, X. Xibin and Y. Yan, "Dynamic resource management in the fourth generation wireless systems," in *Proc. ICCT*, vol. 2, Beijing, China, April 2003, pp. 1095–1098.
- [31] W. Wang, T. Ottosson, M. Sternad, A. Ahlen and A. Svensson, "Impact of multiuser diversity and channel variability on adaptive OFDM," in *Proc. IEEE VTC–Fall*, vol. 1, Orlando, FL, October 2003, pp. 547–551.
- [32] B. Classon, P. Sartori, V. Nangia, Z. Xiangyang and K. Baum, "Multi-dimensional adaptation and Multi-user scheduling techniques for wireless OFDM systems," in *Proc. IEEE ICC*, vol. 3, May 2003, pp. 2251–2255.
- [33] J. Gross, H.-F. Geerdes, H. Karl, A. Wolisz, "Performance analysis of dynamic OFDMA systems with inband signaling," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 427–436, March 2006.
- [34] D. Gesbert and M.S. Alouini, "How Much Feedback is Multi-User Diversity Really Worth," in *Proc. ICC*, vol. 1, June 2004, pp. 234–238.

BIBLIOGRAPHY

- [35] Z.H. Han and Y.H. Lee, "Opportunistic scheduling with partial channel information in OFDMA/FDD systems," in *Proc. IEEE VTC-Fall*, vol. 1, September 2004, pp. 511–514.
- [36] H. Viswanathan and S. Mukherjee, "Performance of Cellular Networks with Relays and Centralized Scheduling," *IEEE Trans. Wireless Comm.*, vol. 4, no. 5, pp. 2318–2328, September 2005.
- [37] H. Hu, H. Yanikomeroglu, D. Falconer and S. Periyalwar, "Range Extension without Capacity Penalty in Cellular Networks with Digital Fixed Relays," in *IEEE GLOBECOM*, Dallas, Texas, December 2004.
- [38] D. Zheng, J. Zhang and J. Sadowsky, "Hierarchical Multiuser Diversity (HMD) Transmission Scheme," in *IEEE GLOBECOM*, San Francisco, CA, December 2003.
- [39] A. Sendonaris, E. Erkip and A. Aazhang, "User Cooperation Diversity—Part I: System Description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, November 2003.
- [40] ———, "User Cooperation Diversity, Part II: Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, November 2003.
- [41] J.N. Laneman and G.W. Wornell, "Distributed space-time-coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Info. Theory*, vol. 49, pp. 2415–2425, November 2003.
- [42] J.C.H. Lin and A. Stefanov, "Coded Cooperation for OFDM Systems," in *Proc. of the IEEE International Conference On Wireless Networks, Communications and Mobile Computing*, Maui, Hawaii, June 2005.
- [43] H. Mheidat and M. Uysal, "Distributed Space-Time Block Coded OFDM for Relay-Assisted Transmission," in *Proc. of the IEEE International Conference On Communications*, Istanbul, Turkey, June 2006.
- [44] M. Herdin, "A Chunk Based OFDM Amplify-and-Forward Relaying Scheme for 4G Mobile Radio Systems," in *Proc. of the IEEE International Conference on Communications*, Istanbul, Turkey, June 2006.
- [45] K. Hayashi et al., "Optimum Relay Position for Differential Amplify-and-Forward Cooperative Communications," in *Proc. of the 9th International Symposium On Wireless Personal Multimedia Communications*, San Diego, CA, September 2006.

BIBLIOGRAPHY

- [46] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [47] M. X. Goemans and D. P. Williamson, “Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming,” *Journal of the Association for Computing Machinery*, vol. 42, pp. 1115–1145, 1995.
- [48] A. Jalali, R. Padovani and R. Pankaj, “Data Throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System,” in *Proc. IEEE VTC*, vol. 3, Tokyo, Japan, May 2000, pp. 1854–1858.
- [49] “Specification of the performance evaluation methodology and the target performance,” IST-2001-32620 MATRICE, D1.3, December 2002.
- [50] S. Yoon et al., “Orthogonal frequency division multiple access with an aggregated sub-channel structure and statistical channel quality measurements,” in *Proc. IEEE VTC*, vol. 2, Los Angeles, CA, September 2004, pp. 1023–1027.
- [51] M. Kaneko, P. Popovski and J. Dahl, “Proportional Fairness in Multi-Carrier System: Upper Bound and Approximation Algorithms,” *IEEE Comm. Letters*, vol. 10, no. 6, pp. 462–464, June 2006.
- [52] G. Senarath et al., “Preliminary Performance Benefit of Single-Hop OFDMA Relay in IEEE 802.16,” IEEE 802.16 MMR, S80216e – 05/010r1, Sept 2005.
- [53] S. Ichitsubo et al., “2 GHz-Band Propagation Loss Prediction in Urban Areas; Antenna Heights Ranging from Ground to Building Roof,” in *IEICE Technical Report, AP 95-15*, May 1996.
- [54] K. Nishimori, R. Di Taranto, H. Yomo, P. Popovski, Y. Takatori, R. Prasad and S. Kubota, “Spatial Opportunity for Cognitive radio Systems with Heterogeneous Path Loss Conditions,” in *IEEE VTC-Spring*, Dulin, Ireland, April 2007, pp. 2631–2635.
- [55] I-K. Fu et al., “Reverse Link Performance of Relay-based Cellular Systems in Manhattan-like Scenario,” IEEE 802.16 MMR, C80216mmr – 06_004r1, Jan 2006.
- [56] M. Newton and J. Thompson, “Classification and Generation of Non-Uniform User Distributions for Cellular Multi-Hop Networks,” in *IEEE ICC*, Istanbul, Turkey, June 2006, pp. 4549–4553.

BIBLIOGRAPHY

- [57] Y. Zhao, R. Adve and T.J. Lim, “Improving Amplify-and-Forward Relay Networks: Optimal Power Allocation versus Selection,” in *Proc. of the IEEE International Symposium On Information Theory*, Seattle, WA, July 2006.
- [58] A. Papoulis and S. Pillai, *Probability, Random Variables and Stochastic Processes*. New York, USA: McGraw Hill Companies, 2002.

Curriculum Vitae

Megumi Kaneko was born in Tokyo, Japan, on May 19, 1981. In June 2004, she jointly received a MSc. degree in Electrical Engineering from Aalborg University, Denmark and a MSc. degree in Information Technology from Institut National des Télécommunications (INT), France. From October 2004, she is a Ph.D. student at the Department of Electronic Systems, Aalborg University. The first two years of her Ph.D. studies were devoted to the research work within the JADE project, a collaborative project between Aalborg University and Samsung Electronics, Korea, for the definition of PHY and MAC layer protocols for the 4G wireless system. From January to July 2007, she was a visiting researcher in Kyoto University, Japan. Her research interests include radio resource allocation and scheduling for Multi-Carrier systems, cross-layer protocol design (MAC/PHY), wireless relaying and cooperation techniques.

